M·H

PCT/AU99/00691

09/786043

AU99/691

4

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Patent Office
Canberra

I, LEANNE MYNOTT, TEAM LEADER PATENT OPERATIONS hereby certify that annexed is a true copy of the Provisional specification in connection with Application No. PP 5512 for a patent by MONASH UNIVERSITY filed on 27 August 1998.

WITNESS my hand this
Twenty-fourth day of September 1999

L. Mynott

LEANNE MYNOTT
TEAM LEADER PATENT OPERATIONS

Regulation 3.2

Monash University

A U S T R A L I A

Patents Act 1990

PROVISIONAL SPECIFICATION

for the invention entitled:

"NOVEL THERAPEUTIC MOLECULES AND USES THEREFOR"

The invention is described in the following statement:

NOVEL THERAPEUTIC MOLECULES AND USES THEREFOR

The present invention relates generally to novel molecules capable of, *inter alia*, controlling cellular functional activity such as proliferation, differentiation and/or transcriptional regulation and to genetic sequences encoding same. More particularly, the present invention relates to novel members of the ETS family of proteins, referred to herein as "ELF5", and to genetic sequences encoding same. The molecules of the present invention are useful, for example, in therapy, diagnosis, antibody generation and as a screening tool for agents capable of modulating transcriptional events during cellular functioning such as in tumorigenesis.

Background of the Invention

The ETS family of transcription factors share a conserved DNA binding domain, termed the 'ETS domain', first identified in the *gag-myb-ets* fusion protein of avian leukemia virus E26 (Nunn *et al.*, 1983; Watson *et al.*, 1988; Karim *et al.*, 1990; Gutman and Wasylyk, 1991; Seth *et al.*, 1992). The ETS domain recognises and binds to purine rich GGA(A/T) core motifs in the promoters and enhancers of various target genes (Macleod *et al.*, 1992; Wasylyk *et al.*, 1993; Janknecht and Nordheim, 1993; Werner *et al.*, 1995; Kodandapani *et al.*, 1996). The ETS family does not maintain overall similarity outside of the ETS domain, but can be grouped into subfamilies based upon variation within the ETS domain, and also by the arrangement and presence of other domains, such as those involved in transactivation and sites of phosphorylation (Lautenberger *et al*, 1992; Wasylyk *et al.*, 1993; Janknecht and Nordheim, 1993). Over 30 ETS gene family members have been identified in species ranging from sea urchin to human.

Many ETS factors have been implicated in the control of cellular proliferation and tumorigenesis (Seth *et al.*, 1992; Macleod *et al.*, 1992; Wasylyk *et al.*, 1993; Janknecht and Nordheim, 1993; Scott *et al.*, 1994a; Muthusamy *et al.*, 1995). *ETS1, ETS2, ERG2* and *PU.1* are proto-oncogenes with mitogenic and transforming activity when

overexpressed in fibroblasts (Seth *et al.*, 1989; Seth and Papas, 1990; Hart *et al.*, 1995; Moreau-Gachelin *et al.*, 1996). In addition, chromosomal translocations involving ETS family members are associated with different human cancers. *ERG* and *ERGB/FLI1* are fused to the *EWS* gene in t(21;22) and t(11;22) translocations, respectively, in Ewing's sarcoma and other primitive neuroectodermal tumors (Sorensen *et al.*, 1994; Ida *et al.*, 1995). *FEV* is fused to *EWS* in a subset of Ewing's tumors in t(2;22) (Peter *et al.*, 1997). *TEL* is fused to the platelet-derived growth factor receptor beta (PDGFRβ) gene in t(5;12) translocations of chronic myelomonocytic leukemia, and to the acute myeloid leukemia 1 (AML1) transcription factor gene in t(12;21) translocations of acute lymphoblastic leukemia (Golub *et al.*, 1994, 1995). Fusion of *TEL* to the receptor-associated kinase JAK2 results in early pre-B acute lymphoid leukemia in t(9;12), and in a typical chronic myelogenous leukemia in t(9;15;12) (Peeters *et al.*, 1997). Expression of *Spil* and *Flil* can be activated by position specific integration of the Friend murine leukemia virus in murine erythroleukemias (Ben-David *et al.*, 1991). Also, ETS1, ETS2 and ERG regulate the expression of metalloproteinase genes, such as stromelysin and collagenase (Buttice and Kurkinen, 1993; Buttice *et al.*, 1996; Wasylyk *et al.*, 1991), which are important for extracellular matrix degradation concomitant with tumor vascularization (angiogenesis) and metastasis.

ETS factors also have important developmental roles. *Pointed P2* and *yan* play critical roles in *Drosophila* eye development (O'Neill *et al.*, 1994). *ETS2* is involved in skeletal/cartilage development (Sumarsono *et al.*, 1996). *PU.1* null mutation results in haematopoietic abnormalities (McKercher *et al.*, 1996), and *ETS1* is involved in transactivation of genes required for T cell function (Muthusamy *et al.*, 1995; Sun *et al.*, 1995; Thomas *et al.*, 1995; Thomas *et al.*, 1997) and angiogenesis (Wasylyk *et al.*, 1991; Vandenbunder *et al.*, 1994; Wernert *et al.*, 1992).

The ETS factors are almost all expressed in haematopoietic lineages (Bhat *et al.*, 1989; Bhat *et al.*, 1990; Kola *et al.*, 1993), and indeed appear to function predominantly in these cells and their related neoplasms. However, the most common solid tumors in humans are

carcinomas which arise from the transformation of epithelial cells. Transformed breast epithelial cells, for example, have been shown to express ETS family members GABPα, PEA3, ELF1, ETS1 and ELK1 (Scott *et al.*, 1994b; Delannoy-Courdent *et al.*, 1996), but expression of these ETS family members is not restricted to epithelial cells. One ETS family member, ELF3/ESX/ESE-1/ERT, has recently emerged with epithelial and epithelial-cancer specific expression (Tymms *et al.*, 1997; Chang *et al.*, 1997; Choi *et al.*, 1998; Oettgen *et al.*, 1997).

In work leading up to the present invention, the inventors have identified and sequenced a novel member of the ETS family, designated herein "ELF5".

Summary of the Invention

Bibliographic details of the publications referred to by author in this specification are collected at the end of the description. Sequence Identity Numbers (SEQ ID NOs.) for the nucleotide and amino acid sequences referred to in the specification are defined following the bibliography. A summary of the SEQ ID NOs. is provided before the Examples.

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

One aspect of the present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding ELF5 wherein said ELF5 comprises an ETS domain.

Another aspect of the present invention provides a nucleic acid molecule comprising a nucleotide sequence encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in one of SEQ ID NO:2 or 4 or a derivative thereof or having at least about 45% or greater similarity to one or more of SEQ ID NO:2 or 4 or a derivative thereof.

Yet another aspect of the present invention provides a nucleic acid molecule comprising a nucleotide sequence encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:7 or having at least about 45% or greater similarity to a sequence comprising the amino acid sequence set forth in SEQ ID NO:7 or a derivative thereof.

Still yet another aspect of the present invention contemplates a nucleic acid molecule comprising a nucleotide sequence substantially as set forth in one of SEQ ID NO:1 or 3 or a derivative thereof capable of hybridising to one of SEQ ID NO:1 or 3 under low stringency conditions at 42°C and which encodes an amino acid sequence corresponding to an amino acid sequence set forth in one of SEQ ID NO:2 or 4 or a sequence having at least about 45% similarity to one or more of SEQ ID NO:2 or 4.

A further aspect of the present invention contemplates a nucleic acid molecule comprising a nucleotide sequence substantially as set forth in one of SEQ ID NO:5 or 6 or a derivative thereof capable of hybridising to one of SEQ ID NO:5 or 6 under low stringency conditions at 42°C and which encodes an amino acid sequence corresponding to an amino acid sequence set forth in SEQ ID NO:7 or a sequence having at least about 45% similarity to SEQ ID NO:7.

Yet another further aspect of the present invention provides a nucleotide sequence corresponding to *ELF5* is a cDNA sequence comprising a sequence of nucleotides as set forth in one of SEQ ID NO:1 or 3 or is a derivative thereof including a nucleotide sequence having similarity to one of SEQ ID NO:1 or 3 and which encodes an amino acid sequence corresponding to an amino acid sequence set forth in one of SEQ ID NO:2 or 4 or a sequence having at least about 45% similarity to one or more of SEQ ID NO:2 or 4.

Still yet another further aspect of the present invention provides a nucleotide sequence corresponding to *ELF5* is a cDNA sequence comprising a sequence of nucleotides as set forth in one of SEQ ID NO:5 or 6 or is a derivative thereof including a nucleotide sequence having similarity to one of SEQ ID NO:5 or 6 and which encodes an amino acid sequence

corresponding to an amino acid sequence as set forth in SEQ ID NO:7 or a sequence having at least about a 45% similarity to SEQ ID NO:7.

Another aspect of the present invention contemplates a method of modulating activity of ELF5 in a mammal, said method comprising administering to said mammal a modulating effective amount of an agent for a time and under conditions sufficient to increase or decrease ELF5 activity.

Yet another aspect of the present invention contemplates a method of modulating cellular functional activity in a mammal said method comprising administering to said mammal an effective amount of an agent for a time and under conditions sufficient to modulate the expression of a nucleotide sequence encoding *ELF5* or sufficient to modulate the activity of ELF5.

Still yet another aspect of the present invention contemplates a method of modulating cellular functional activity in a mammal said method comprising administering to said mammal an effective amount of ELF5 or *ELF5*.

A further aspect of the present invention relates to a method of treating a mammal said method comprising administering to said mammal an effective amount of an agent for a time and under conditions sufficient to modulate the expression of *ELF5* or sufficient to modulate the activity of ELF5 wherein said modulation results in modulation of cellular functional activity.

Yet another further aspect of the present invention relates to a method of treating a mammal said method comprising administering to said mammal an effective amount of ELF5 or *ELF5* for a time and under conditions sufficient to modulate cellular functional activity.

Still yet another further aspect of the present invention relates to the use of an agent capable of modulating the expression of *ELF5* or modulating the activity of ELF5 in the manufacture of a medicament for the modulation of cellular functional activity.

Another aspect of the present invention relates to the use of ELF5 or *ELF5* in the manufacture of a medicament for the modulation of cellular functional activity.

Yet another aspect of the present invention relates to agents for use in modulating *ELF5* expression or ELF5 activity wherein said modulation results in modulation of cellular functional activity.

Still yet another aspect of the present invention relates to ELF5 or *ELF5* for use in modulating cellular functional activity.

A further aspect of the present invention contemplates a pharmaceutical composition comprising *ELF5*, ELF5 or an agent capable of modulating *ELF5* expression or ELF5 activity together with one or more pharmaceutically acceptable carriers and/or diluents. *ELF5*, ELF5 or said agent are referred to as the active ingredients.

Another further aspect of the present invention contemplates a method for detecting ELF5 or *ELF5* mRNA in a biological sample from a subject said method comprising contacting said biological sample with an antibody specific for ELF5 or *ELF5* mRNA or its derivatives or homologs for a time and under conditions sufficient for an antibody-ELF5 or antibody-ELF5 mRNA complex to form, and then detecting said complex.

Single and three letter abbreviations used throughout the specification are defined in Table 1.

TABLE 1

Single and three letter amino acid abbreviations

Amino Acid	Three-letter Abbreviation	One-letter Symbol
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V
Any residue	Xaa	X

Brief Description of the Drawings

Figure 1 is a schematic representation of murine *ELF5* cDNA sequence and relationship to mRNA transcripts. (a) The nucleotide sequence of murine *ELF5* is shown. Breaks in the sequence indicate the source of sequence data; the central region (92-1528) was sequenced from lambda clones, and 5 prime and 3 prime were added from sequencing of RACE PCR products. Numbering of the nucleotides, starting with the most 5-prime sequences obtained, are indicated on the right. The open reading frame (ORF) is shown in capital letters, with the initiating start and stop codons underlined. A stop codon, in the same reading frame as the ORF, but 5 prime to the initiating codon, is also underlined. The ETS domain is indicated in a shaded box. Putative polyadenylation signals are underlined with dashed lines. A/T rich tracts in the 3 prime untranslated region are boxed. (b) Northern blot analysis of day 14 mouse placenta: lane 1, probed with random-prime-labeled 940 bp *Sty1* murine *ELF5* cDNA fragment (probe 1); lane 2, probed with random-prime-labeled murine *ELF5* 696 bp 3'-RACE PCR product (probe 2). Positions of 28S and 18S markers are indicated. Both lanes were also probed with GAPDH cDNA (lower panels).

Figure 2 is a schematic representation of (a) Comparison of human and mouse ORFs. Amino acid sequences present in both human and mouse ELF5 are shaded. The ETS domain is boxed with a solid line and the pointed domain with a dashed line. Putative phosphorylation sites, conserved between the two species are circled and labeled as CKII (casein kinase II), PKC (protein kinase C) or TyP (tyrosine kinase) substrates. (b) Comparison of the ETS domain of human and mouse ELF5 with those of known members of the ETS gene family. The alignment was generated using CLUSTAL W (Thompson *et al.*, 1994) with the default settings, and the result was subsequently adjusted manually. The ETS factors examined are labeled on the left and include hELF3, mELF3, hNERF, dETS4, dE74A, hELF1, hELK1, hTEL, hERM, mER81, mPEA3, mGABP∝ mERP, dETS6, mPU1, hPE1, hSAP1, hSPIB, dYAN, hERG, mFLI1, dELG, dETS3, mETS1, mETS2, mER71, where 'h' denotes human, 'm' mouse and 'd' *Drosophila*. The ETS consensus sequence is a list of the amino acids most often conserved between ETS family members. Shading denotes amino acid identity with human ELF5, and the percent identity of each ETS domain is

indicated on the right. (c) Phylogenetic tree of the ETS domain produced by maximum likelihood analysis. The alignment in Figure 2b was analysed using the JTT-F substitution model (Jones *et al.*, 1992) and local bootstrap values were estimated for all internal branches, both by using PROTML in Q mode followed by a second run in R mode (Adachi and Hasegawa, 1996). An underlying assumption of the phylogenetic analysis is that the amino acid content does not vary significantly among the sequences. This assumption was not assessed because tools for doing so are still under development (LSJ, unpublished work). Therefore, the tree may be the result of both historical and compositional components. The four points at which gene duplications have been inferred are marked A, B, C and D. (d) Comparison of the pointed domain of human and mouse ELF5 with those of other members of the ETS family. The ETS factors examined are labeled on the left and include hERG, hELF3, hTEL, hGABP∝, hETS1, hETS2, dYAN and dPOINTEDP2. Other labels and conventions are as described for Figure 2b.

Figure 3 is a schematic representation of the chromosomal localization of human *ELF5*. Human chromosomal localization of ELF5 was performed by PCR using gene specific primers and the Genebridge 4 Radiation Hybrid DNA panel (UK HGMP Resource Centre). Diagram based upon PCR results (data not shown) showing localization of *ELF5* within chromosome one, with respect to adjacent marker obtained from mapping data (see text).

Figure 4 is a photographic representation of *ELF5* expression in mouse tissues. Positions of 28S and 18S markers are indicated. *ELF5a* and *ELF5b* transcripts are indicated. (a) Northern analysis of adult mouse tissues probed with murine *ELF5* cDNA (top panels) and GAPDH cDNA (lower panels). Abbreviations; Li: liver; Lu: lung; Br: brain; Ki: kidney; He: heart; Sm: small intestine; Sp: spleen; Th: thymus; St: stomach; Ov: ovary; Pa: pancreas; To: tongue; Sk: skeletal muscle; Bl: bladder; 2Fa: day 2 pregnant fat; 2 Ma: day 2 pregnant mammary gland; 10 Fa: day 10 pregnant fat; 10 Ma: day 10 pregnant mammary gland; Co: colon. Arrow indicates position of brain specific transcript (see text). (b) Northern analysis as above, but using RNA from day 1 neonate mouse tissues. Additional abbreviation; In: intestine. Arrow indicates position of large transcript (see text). (c) Northern analysis as above, but using RNA from day 16, 17 and 19 embryonic tissues. (d) Northern analysis as

above, but using RNA from day 9.5 to day 19 placental tissues as indicated.

Figure 5 is a photographic representation of *ELF5* expression in human tissues and cell lines. (a) Northern analysis of adult human tissues probed with human *ELF5* cDNA (top panels) and β-Actin cDNA (lower panels). The single *ELF5* transcript is indicated. Other labels and conventions are as for Figure 4. Abbreviations; He: heart; Br: brain; Pl: placenta; Lu: lung; Li: liver; Sk: skeletal muscle; Ki: kidney; Pa: pancreas; Sp: spleen; Th: thymus; Pr: prostate; Te: testis; Ov: ovary; Sm: small intestine; Co: colon mucosa; PBL: peripheral blood lymphocytes. (b) RNAse protection analysis of *ELF5* and GAPDH in cell lines; 1: CaOv-3 (ovarian carcinoma); 2: BT-549 (ductal breast carcinoma); 3: ZR-75-1 (breast carcinoma); 4: T47D (ductal breast carcinoma, progesterone sensitive); 5: 786-O (renal adenocarcinoma); 6: SK-HEP-1 (liver adenocarcinoma); 7: A549 (lung adenocarcinoma); 8: CCL32SK (primary fibroblast); 9: MEL28 (melanoma); 10: WISH (amnion carcinoma); 11: Jurkat (T cell leukemia); 12: DU145 (prostate carcinoma); 13: PC3 (prostate carcinoma); 14: HEC-1 (endometrium carcinoma); 15: K562 (erythroid leukemia). (c) Southern analysis of *ELF5* in *Bgl*II digested genomic DNA from cell lines; 1: normal blood; 2: BT-549 (ductal breast carcinoma); 3: ZR-75-1 (breast carcinoma); 4: T47D (ductal breast carcinoma); 5: NCI-H1299 (large cell lung carcinoma); 6: NCI-HI87 (small cell lung carcinoma); 7: NCI-H322 (bronchioalveolar carcinoma); 8: NCI-H358 (bronchioalveolar carcinoma); 9: NCI-H522 (lung adenocarcinoma); 10: SK-LU-1 (lung adenocarcinoma); 11: NCI-H441 (bronchioalveolar carcinoma); 12: NCI-H460 (large cell lung carcinoma); 13: NCI-H661 (large cell lung carcinoma).

Figure 6 is a photographic representation of ELF5 binding to consensus ETS binding sequences. (a) His-tagged ELF5 recombinant protein, present in *E. coli* lysates (lane 2), was purified by metal-affinity chromatography to approximately 90% (lane 3) and eluted with imadazole (lane 4). (b) Specific DNA binding of Elf5 was analysed by electrophoretic mobility shift assay (EMSA), using labeled double-stranded oligonucleotides as probes. E74 contains a consensus binding site for ETS family members (lane 1). E74ml is a mutant oligonucleotide based on E74, but with the core GGAA replaced by AGAA (lane 2). Binding to other consensus ETS sites was analysed by the ability of a 100-fold excess of unlabeled

double-stranded oligonucleotide to compete with E74 for Elf5 binding GMETS contains an ETS binding site from the human GM-CSF promoter (lane 6). ERBB2 contains an ETS binding site from the human *erbB2/*HER2 promoter (lane 7). MSV contains an ETS binding site present in the long terminal repeat of the Moloney sarcoma virus (lane 8). AP1 contains a consensus AP1 binding site used as a negative control ELF5-DNA complexes are marked. Binding of ETS1 to E74 was used as a positive control (lane 10).

Figure 7 is a graphical representation of transactivation by ELF5. COS cells were co-transfected with CAT reporter and Elf5 expression constructs. Transcription of the CAT gene was driven by the thymidine kinase (tk) minimal promoter with five copies of the polyomavirus enhancer inserted upstream (p5Xpoly). The polyomavirus enhancer contains adjacent ETS and AP1 binding sites. The ELF5 sense construct (pBOSElf5as) was designed to express ELF5 protein, and the *ELF5* anti-s construct (pBOSElf5as) to produce anti-sense transcripts. In the absence of expression construct the equivalent amount of base vector (pEFBOS) was co-transfected. COS cells were processed for CAT assays and the results of at least four replicates are shown as the mean with standard error of the mean (s.e.m.) bars. Statistically significant results are indicated by asterisks. A single asterisk indicates moderate significance ($0.05>P>0.01$) and triple asterisks indicate very high significance ($P<0.001$).

Figure 8 is a photographic representation of breast tissue sections from parafin-embedded samples which had been hybridized with ELF5 antisense RNA.

Detailed Description of the Invention

The present invention is predicated, in part, on the identification of a novel member of the ETS family of molecules, termed ELF5. The identification of this novel molecule permits the identification and rational design of a range of products for use in therapy, diagnosis and antibody generation involving, for example, regulation of cellular functional activity such as cellular proliferation. These therapeutic molecules may also act as either antagonists or agonists of ELF5 function and will be useful, *inter alia*, in cancer and autoimmune disease therapy.

Accordingly, one aspect of the present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding ELF5 wherein said ELF5 comprises an ETS domain.

Reference to an "ETS domain" should be understood as a reference to a protein domain which recognises and binds to a purine rich GGA(A/T) motif of a promoter or enhancer (Macleod *et al.*, 1992; Wasylyk *et al.*, 1993; Janknecht and Nordheim, 1993; Werner *et al.*, 1995; Kodandapani *et al.*, 1996). The ETS domain may be continuous, meaning that it is comprised of a continuous sequence of amino acids, or it may be discontinuous, meaning that it is comprised of individual amino acids or sequences of amino acids from two or more separate regions of the protein and which are brought into proximity with one another to form the ETS domain due to the secondary, tertiary or quaternary structure of the protein.

More particularly, the present invention provides a nucleic acid molecule comprising a nucleotide sequence encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in one of SEQ ID NO:2 or 4 or a derivative thereof or having at least about 45% or greater similarity to one or more of SEQ ID NO:2 or 4 or a derivative thereof.

Another aspect of the present invention provides a nucleic acid molecule comprising a nucleotide sequence encoding or complementary to a sequence encoding an amino acid

sequence substantially as set forth in SEQ ID NO:7 or having at least about 45% or greater similarity to a sequence comprising the amino acid sequence set forth in SEQ ID NO:7 or a derivative thereof.

The term "similarity" as used herein includes exact identity between compared sequences at the nucleotide or amino acid levels. Where there is non-identity of the nucleotide level "similarity" includes differences between sequences which result in different amino acids that are nevertheless related to each other at the structural, functional, biochemical and/or conformational levels. Where there is non-identity at the amino acid level, "similarity" includes amino acids that are nevertheless related to each other at the structural, functional, biochemical and/or conformational levels. The percentage similarity may be greater than 50% such as at least 70% or at least 80% or at least 90% or higher.

Another aspect of the present invention contemplates a nucleic acid molecule comprising a nucleotide sequence substantially as set forth in one of SEQ ID NO:1 or 3 or a derivative thereof capable of hybridising to one of SEQ ID NO:1 or 3 under low stringency conditions at 42°C and which encodes an amino acid sequence corresponding to an amino acid sequence set forth in one of SEQ ID NO:2 or 4 or a sequence having at least about 45% similarity to one or more of SEQ ID NO:2 or 4.

More particularly, the present invention contemplates a nucleic acid molecule comprising a sequence of nucleotides substantially as set forth in SEQ ID NO:1 or 3.

Another aspect of the present invention contemplates a nucleic acid molecule comprising a nucleotide sequence substantially as set forth in one of SEQ ID NO:5 or 6 or a derivative thereof capable of hybridising to one of SEQ ID NO:5 or 6 under low stringency conditions at 42°C and which encodes an amino acid sequence corresponding to an amino acid sequence set forth in SEQ ID NO:7 or a sequence having at least about 45% similarity to SEQ ID NO:7.

More particularly, the present invention contemplates a nucleic acid molecule comprising a

sequence of nucleotides substantially as set forth in SEQ ID NO:5 or 6.

Reference herein to a low stringency at 42°C includes and encompasses from at least about 1% v/v to at least about 15% v/v formamide and from at least about 1M to at least about 2M salt for hybridisation, and at least about 1M to at least about 2M salt for washing conditions. Alternative stringency conditions may be applied where necessary, such as medium stringency, which includes and encompasses from at least about 16% v/v to at least about 30% v/v formamide and from at least about 0.5M to at least about 0.9M salt for hybridisation, and at least about 0.5M to at least about 0.9M salt for washing conditions, or high stringency, which includes and encompasses from at least about 31% v/v to at least about 50% v/v formamide and from at least about 0.01M to at least about 0.15M salt for hybridisation, and at least about 0.01M to at least about 0.15M salt for washing conditions. In general, washing is carried out at $T_m = 69.3 + 0.41\ (G + C)\ \%$ [19] = -12°C. However, the T_m of a duplex DNA decreases by 1°C with every increase of 1% in the number of mismatched based pairs (20).

The nucleic acid molecule according to this aspect of the present invention corresponds herein to "*ELF5*". This gene has been determined in accordance with the present invention to encode a protein which displays specific binding to DNA sequences comprising a GGA(A/T) core. The product of the *ELF5* gene is referred to herein as ELF5. ELF5 is defined as belonging to the ETS family of transcription factors due to its expression of an ETS domain which recognises and binds the purine rich GGA(A/T) core motifs. ELF5 is a protein for which splice variants exist, thereby resulting in the expression of a variety of isoforms. Human ELF5 and human ELF5 short transcript are examples of 2 isoforms which differ in size due to the splicing out of exon regions from the *ELF5* mRNA molecule encoding the ELF5 short transcript. Murine *ELF5a* and *ELF5b* are examples of 2 mRNA transcripts which differ in the length of the 3' untranslated region. Human ELF5 and ELF5 short transcript are defined by the amino acid sequences set forth in SEQ ID NO: 2 and 4, respectively and murine ELF5 is defined by the amino acid sequence set forth in SEQ ID NO:7. The cDNA nucleotide sequences for human *ELF5* and *ELF5* short transcript are defined by the nucleotide sequences set forth in SEQ ID NO:1 and 3, respectively, and

murine *ELF5a* and *ELFb* are defined by the nucleotide sequences set forth in SEQ ID NO:5 and 6, respectively.

The nucleic acid molecules encoding ELF5 are preferably a sequence of deoxyribonucleic acids such as cDNA sequences or genomic sequences. A genomic sequence may also comprise exons and introns. A genomic sequence may also include a promoter region or other regulatory region.

Reference hereinafter to "ELF5" and "*ELF5*" should be understood as a reference to all forms of ELF5 and *ELF5*, respectively, including by way of example the two mRNA transcripts, *ELF5a* and *ELF5b*, observed in the mouse. Without limiting the invention in any way, sequence analysis of murine ELF5 has revealed two discrete polyadenylation signals present in the 3' untranslated region (UTR). The first of these (at 1391 bp) appears to be an overlapping poly(A)$^+$ recognition signal AATTAA and ATTAAAA. The second is a consensus polyadenylation signal, AATAAA, at 2181 bp. Sequence analysis of human *ELF5* has also revealed two mRNA transcripts arising from the splicing out of part of the exon region. Accordingly, the present invention should be understood to extend to all cDNA and peptide isoforms arising from alternative splicing of ELF5 mRNA.

The protein and/or gene is preferably from a human, primate, livestock animal (eg. sheep, pig, cow, horse, donkey) laboratory test animal (eg. mouse, rat, rabbit, guinea pig) companion animal (eg. dog, cat), captive wild animal (eg. fox, kangaroo, deer), aves (eg. chicken, geese, duck, emu, ostrich), reptile or fish.

The term "protein" should be understood to encompass peptides, polypeptides and proteins. The protein may be glycosylated or unglycosylated, phosphorylated or dephosphorylated to various degrees and/or may contain a range of other molecules fused, linked, bound or otherwise associated to the protein such as amino acids, lipids, carbohydrates or other peptides, polypeptides or proteins. Reference hereinafer to a "protein" includes a protein comprising a sequence of amino acids as well as a protein associated with other molecules such as amino acids, lipids, carbohydrates or other peptides, polypeptides or proteins.

Derivatives include fragments, parts, portions, chemical equivalents, mutants, homologs, mimetics from natural, synthetic or recombinant sources including fusion proteins. Derivatives may be derived from insertion, deletion or substitution of amino acids. Amino acid insertional derivatives include amino and/or carboxylic terminal fusions as well as intrasequence insertions of single or multiple amino acids. Insertional amino acid sequence variants are those in which one or more amino acid residues are introduced into a predetermined site in the protein although random insertion is also possible with suitable screening of the resulting product. Deletional variants are characterized by the removal of one or more amino acids from the sequence. Substitutional amino acid variants are those in which at least one residue in the sequence has been removed and a different residue inserted in its place. Additions to amino acid sequences including fusions with other peptides, polypeptides or proteins.

The derivatives of ELF5 include fragments having particular epitopes or parts of the entire ELF5 protein fused to peptides, polypeptides or other proteinaceous or non-proteinaceous molecules. For example, ELF5 or derivative thereof may be fused to a molecule to facilitate its entry into a cell. Analogs of ELF5 contemplated herein include, but are not limited to, modification to side chains, incorporating of unnatural amino acids and/or their derivatives during peptide, polypeptide or protein synthesis and the use of crosslinkers and other methods which impose conformational constraints on the proteinaceous molecules or their analogs. Derivatives of nucleic acid sequences may similarly be derived from single or multiple nucleotide substitutions, deletions and/or additions including fusion with other nucleic acid molecules. The derivatives of the nucleic acid molecules of the present invention include oligonucleotides, PCR primers, antisense molecules, molecules suitable for use in cosuppression and fusion of nucleic acid molecules.

Examples of side chain modifications contemplated by the present invention include modifications of amino groups such as by reductive alkylation by reaction with an aldehyde followed by reduction with $NaBH_4$; amidination with methylacetimidate; acylation with acetic anhydride; carbamoylation of amino groups with cyanate;

trinitrobenzylation of amino groups with 2, 4, 6-trinitrobenzene sulphonic acid (TNBS); acylation of amino groups with succinic anhydride and tetrahydrophthalic anhydride; and pyridoxylation of lysine with pyridoxal-5-phosphate followed by reduction with $NaBH_4$.

The guanidine group of arginine residues may be modified by the formation of heterocyclic condensation products with reagents such as 2,3-butanedione, phenylglyoxal and glyoxal.

The carboxyl group may be modified by carbodiimide activation *via* O-acylisourea formation followed by subsequent derivitisation, for example, to a corresponding amide.

Sulphydryl groups may be modified by methods such as carboxymethylation with iodoacetic acid or iodoacetamide; performic acid oxidation to cysteic acid; formation of a mixed disulphides with other thiol compounds; reaction with maleimide, maleic anhydride or other substituted maleimide; formation of mercurial derivatives using 4-chloromercuribenzoate, 4-chloromercuriphenylsulphonic acid, phenylmercury chloride, 2-chloromercuri-4-nitrophenol and other mercurials; carbamoylation with cyanate at alkaline pH.

Tryptophan residues may be modified by, for example, oxidation with N-bromosuccinimide or alkylation of the indole ring with 2-hydroxy-5-nitrobenzyl bromide or sulphenyl halides. Tyrosine residues on the other hand, may be altered by nitration with tetranitromethane to form a 3-nitrotyrosine derivative.

Modification of the imidazole ring of a histidine residue may be accomplished by alkylation with iodoacetic acid derivatives or N-carboethoxylation with diethylpyrocarbonate.

Examples of incorporating unnatural amino acids and derivatives during protein synthesis include, but are not limited to, use of norleucine, 4-amino butyric acid, 4-amino-3-

hydroxy-5-phenylpentanoic acid, 6-aminohexanoic acid, t-butylglycine, norvaline, phenylglycine, ornithine, sarcosine, 4-amino-3-hydroxy-6-methylheptanoic acid, 2-thienyl alanine and/or D-isomers of amino acids. A list of unnatural amino acid contemplated herein is shown in Table 2.

5

TABLE 2

Non-conventional amino acid	Code	Non-conventional amino acid	Code
α-aminobutyric acid	Abu	L-N-methylalanine	Nmala
α-amino-α-methylbutyrate	Mgabu	L-N-methylarginine	Nmarg
aminocyclopropane-carboxylate	Cpro	L-N-methylasparagine	Nmasn
		L-N-methylaspartic acid	Nmasp
aminoisobutyric acid	Aib	L-N-methylcysteine	Nmcys
aminonorbornyl-carboxylate	Norb	L-N-methylglutamine	Nmgln
		L-N-methylglutamic acid	Nmglu
cyclohexylalanine		Chexa L-N-methylhistidine	Nmhis
cyclopentylalanine	Cpen	L-N-methylisolleucine	Nmile
D-alanine	Dal	L-N-methylleucine	Nmleu
D-arginine	Darg	L-N-methyllysine	Nmlys
D-aspartic acid	Dasp	L-N-methylmethionine	Nmmet
D-cysteine	Dcys	L-N-methylnorleucine	Nmnle
D-glutamine	Dgln	L-N-methylnorvaline	Nmnva
D-glutamic acid	Dglu	L-N-methylornithine	Nmorn
D-histidine	Dhis	L-N-methylphenylalanine	Nmphe
D-isoleucine	Dile	L-N-methylproline	Nmpro
D-leucine	Dleu	L-N-methylserine	Nmser
D-lysine	Dlys	L-N-methylthreonine	Nmthr
D-methionine	Dmet	L-N-methyltryptophan	Nmtrp
D-ornithine	Dorn	L-N-methyltyrosine	Nmtyr
D-phenylalanine	Dphe	L-N-methylvaline	Nmval
D-proline	Dpro	L-N-methylethylglycine	Nmetg
D-serine	Dser	L-N-methyl-t-butylglycine	Nmtbug
D-threonine	Dthr	L-norleucine	Nle

	D-tryptophan	Dtrp	L-norvaline	Nva
	D-tyrosine	Dtyr	α-methyl-aminoisobutyrate	Maib
	D-valine	Dval	α-methyl-γ-aminobutyrate	Mgabu
	D-α-methylalanine	Dmala	α-methylcyclohexylalanine	Mchexa
5	D-α-methylarginine	Dmarg	α-methylcylcopentylalanine	Mcpen
	D-α-methylasparagine	Dmasn	α-methyl-α-napthylalanine	Manap
	D-α-methylaspartate	Dmasp	α-methylpenicillamine	Mpen
	D-α-methylcysteine	Dmcys	N-(4-aminobutyl)glycine	Nglu
	D-α-methylglutamine	Dmgln	N-(2-aminoethyl)glycine	Naeg
10	D-α-methylhistidine	Dmhis	N-(3-aminopropyl)glycine	Norn
	D-α-methylisoleucine	Dmile	N-amino-α-methylbutyrate	Nmaabu
	D-α-methylleucine	Dmleu	α-napthylalanine	Anap
	D-α-methyllysine	Dmlys	N-benzylglycine	Nphe
	D-α-methylmethionine	Dmmet	N-(2-carbamylethyl)glycine	Ngln
15	D-α-methylornithine	Dmorn	N-(carbamylmethyl)glycine	Nasn
	D-α-methylphenylalanine	Dmphe	N-(2-carboxyethyl)glycine	Nglu
	D-α-methylproline	Dmpro	N-(carboxymethyl)glycine	Nasp
	D-α-methylserine	Dmser	N-cyclobutylglycine	Ncbut
	D-α-methylthreonine	Dmthr	N-cycloheptylglycine	Nchep
20	D-α-methyltryptophan	Dmtrp	N-cyclohexylglycine	Nchex
	D-α-methyltyrosine	Dmty	N-cyclodecylglycine	Ncdec
	D-α-methylvaline	Dmval	N-cylcododecylglycine	Ncdod
	D-N-methylalanine	Dnmala	N-cyclooctylglycine	Ncoct
	D-N-methylarginine	Dnmarg	N-cyclopropylglycine	Ncpro
25	D-N-methylasparagine	Dnmasn	N-cycloundecylglycine	Ncund
	D-N-methylaspartate	Dnmasp	N-(2,2-diphenylethyl)glycine	Nbhm
	D-N-methylcysteine	Dnmcys	N-(3,3-diphenylpropyl)glycine	Nbhe
	D-N-methylglutamine	Dnmgln	N-(3-guanidinopropyl)glycine	Narg
	D-N-methylglutamate	Dnmglu	N-(1-hydroxyethyl)glycine	Nthr
30	D-N-methylhistidine	Dnmhis	N-(hydroxyethyl))glycine	Nser

	D-N-methylisoleucine	Dnmile	N-(imidazolylethyl))glycine	Nhis
	D-N-methylleucine	Dnmleu	N-(3-indolylyethyl)glycine	Nhtrp
	D-N-methyllysine	Dnmlys	N-methyl-γ-aminobutyrate	Nmgabu
	N-methylcyclohexylalanine	Nmchexa	D-N-methylmethionine	Dnmmet
	D-N-methylornithine	Dnmorn	N-methylcyclopentylalanine	Nmcpen
	N-methylglycine	Nala	D-N-methylphenylalanine	Dnmphe
	N-methylaminoisobutyrate	Nmaib	D-N-methylproline	Dnmpro
	N-(1-methylpropyl)glycine	Nile	D-N-methylserine	Dnmser
	N-(2-methylpropyl)glycine	Nleu	D-N-methylthreonine	Dnmthr
	D-N-methyltryptophan	Dnmtrp	N-(1-methylethyl)glycine	Nval
	D-N-methyltyrosine	Dnmtyr	N-methyla-napthylalanine	Nmanap
	D-N-methylvaline	Dnmval	N-methylpenicillamine	Nmpen
	γ-aminobutyric acid	Gabu	N-(*p*-hydroxyphenyl)glycine	Nhtyr
	L-*t*-butylglycine	Tbug	N-(thiomethyl)glycine	Ncys
	L-ethylglycine	Etg	penicillamine	Pen
	L-homophenylalanine	Hphe	L-α-methylalanine	Mala
	L-α-methylarginine	Marg	L-α-methylasparagine	Masn
	L-α-methylaspartate	Masp	L-α-methyl-*t*-butylglycine	Mtbug
	L-α-methylcysteine	Mcys	L-methylethylglycine	Metg
	L-α-methylglutamine	Mgln	L-α-methylglutamate	Mglu
	L-α-methylhistidine	Mhis	L-α-methylhomophenylalanine	Mhphe
	L-α-methylisoleucine	Mile	N-(2-methylthioethyl)glycine	Nmet
	L-α-methylleucine	Mleu	L-α-methyllysine	Mlys
	L-α-methylmethionine	Mmet	L-α-methylnorleucine	Mnle
	L-α-methylnorvaline	Mnva	L-α-methylornithine	Morn
	L-α-methylphenylalanine	Mphe	L-α-methylproline	Mpro
	L-α-methylserine	Mser	L-α-methylthreonine	Mthr
	L-α-methyltryptophan	Mtrp	L-α-methyltyrosine	Mtyr
	L-α-methylvaline	Mval	L-N-methylhomophenylalanine	Nmhphe
	N-(N-(2,2-diphenylethyl)	Nnbhm	N-(N-(3,3-diphenylpropyl)	Nnbhe

carbamylmethyl)glycine		carbamylmethyl)glycine
1-carboxy-1-(2,2-diphenyl-ethylamino)cyclopropane	Nmbc	

Crosslinkers can be used, for example, to stabilise 3D conformations, using homo-bifunctional crosslinkers such as the bifunctional imido esters having $(CH_2)_n$ spacer groups with n=1 to n=6, glutaraldehyde, N-hydroxysuccinimide esters and hetero-bifunctional reagents which usually contain an amino-reactive moiety such as N-hydroxysuccinimide and another group specific-reactive moiety such as maleimido or dithio moiety (SH) or carbodiimide (COOH). In addition, peptides can be conformationally constrained by, for example, incorporation of C_α and N_α-methylamino acids, introduction of double bonds between C_α and C_β atoms of amino acids and the formation of cyclic peptides or analogues by introducing covalent bonds such as forming an amide bond between the N and C termini, between two side chains or between a side chain and the N or C terminus.

The nucleic acid molecule of the present invention is preferably in isolated form or ligated to a vector, such as an expression vector. By "isolated" is meant a nucleic acid molecule having undergone at least one purification step and this is conveniently defined, for example, by a composition comprising at least about 10% subject nucleic acid molecule, preferably at least about 20%, more preferably at least about 30%, still more preferably at least about 40-50%, even still more preferably at least about 60-70%, yet even still more preferably 80-90% or greater of subject nucleic acid molecule relative to other components as determined by molecular weight, encoding activity, nucleotide sequence, base composition or other convenient means. The nucleic acid molecule of the present invention may also be considered, in a preferred embodiment, to be biologically pure.

In a particularly preferred embodiment, the nucleotide sequence corresponding to *ELF5* is a cDNA sequence comprising a sequence of nucleotides as set forth in one of SEQ ID NO:1 or 3 or is a derivative thereof including a nucleotide sequence having similarity to one of SEQ ID NO:1 or 3 and which encodes an amino acid sequence corresponding to an amino acid sequence set forth in one of SEQ ID NO:2 or 4 or a sequence having at least about 45% similarity to one or more of SEQ ID NO:2 or 4.

In another particularly preferred embodiment, the nucleotide sequence corresponding to *ELF5* is a cDNA sequence comprising a sequence of nucleotides as set forth in one of SEQ ID NO:5 or 6 or is a derivative thereof including a nucleotide sequence having similarity to one of SEQ ID NO:5 or 6 and which encodes an amino acid sequence corresponding to an amino acid sequence as set forth in SEQ ID NO:7 or a sequence having at least about a 45% similarity to SEQ ID NO:7.

A derivative of the nucleic acid molecule of the present invention also includes nucleic acid molecules capable of hybridising to the nucleotide sequences as set forth in one of SEQ ID NO:1 or 3 or SEQ ID NO:5 or 6 under low stringency conditions. Preferably said low stringency is at 42°C.

The nucleic acid molecule may be ligated to an expression vector capable of expression in a prokaryotic cell (eg. *E. coli*) or a eukaryotic cell (eg. yeast cells, fungal cells, insect cells, mammalian cells or plant cells). The nucleic acid molecule may be ligated or fused or otherwise associated with a nucleic acid molecule encoding another entity such as, for example, a signal peptide, a cytokine or other member of the ETS family.

The expression product is ELF5 having an amino acid sequence set forth in one of SEQ ID NO:2 or 4 or SEQ ID NO:7 or is a derivative or homologue as hereinbefore defined or is a mammalian homologue having an amino acid sequence of at least about 45% similarity to the amino acid sequence set forth in one of SEQ ID NO:2 or 4 or SEQ ID NO:7 or derivative or homologue thereof.

The ELF5 of the present invention may be in multimeric form meaning that two or more molecules are associated together. Where the same ELF5 molecules are associated together, the complex is a homomultimer. An example of a homomultimer is a homodimer. Where at least one ELF5 is associated with at least one non-ELF5 molecule, then the complex is a heteromultimer such as a heterodimer. A heteromultimer may include a molecule of another member of the ETS family or other molecule capable of modulating transcription.

In accordance with the present invention, it is proposed that ELF5 is a molecule which regulates cellular functional activity. Reference to cellular "functional activity" should be understood as a reference to the functions which a cell is capable of performing such as, but in no way limited to, one or more of proliferation, differentiation, cell surface molecule expression, antigen presentation, maintenance of viability, apoptosis, metabolism, signal transduction and molecular mechanisms such as transcription and translation. Without limiting this invention to any one theory or mode of action, human *ELF5* has been mapped to human chromosome 1p36.31 which is a region that frequently undergoes loss of heterozygocity in several types of carcinoma, including breast and colon carcinomas. The expression pattern of *ELF5* and ELF5 in normal and diseased tissues also supports a role for these molecules in the regulation of cellular functional activity and, in particular, in the direct or indirect regulation of tumorigenesis. Even more particularly, it is proposed that ELF5 functions as a transcription factor.

The cloning and sequencing of this gene and its expression product now provides an additional gene for use in the prophylactic and therapeutic treatment of diseases such as those involving aberrant cellular functional activity such as aberrant cellular proliferation. Examples of diseases involving aberrant cellular proliferation include diseases caused by excessive cellular proliferation, such as in tumorigenesis, or diseases caused by inadequate cellular proliferation. Accordingly, the present invention contemplates therapeutic and prophylactic uses of ELF5 amino acid and nucleic acid molecules, in addition to ELF5 agonistic and antagonistic agents, for the regulation of cellular functional activity, such as for example, regulation of proliferation, differentiation and/or regulation of gene expression

by transcriptional regulation.

The present invention contemplates, therefore, a method for modulating expression of *ELF5* in a subject, said method comprising contacting *ELF5* gene with an effective amount of an agent for a time and under conditions sufficient to up-regulate or down-regulate or otherwise modulate expression of *ELF5*. For example, *ELF5* antisense sequences such as oligonucleotides may be introduced into a cell to down-regulate one or more specific functional activities of that cell. Conversely, a nucleic acid molecule encoding ELF5 or a derivative thereof may be introduced to up-regulate one or more specific functional activities of any cell not expressing the endogenous *ELF5* gene.

Another aspect of the present invention contemplates a method of modulating activity of ELF5 in a mammal, said method comprising administering to said mammal a modulating effective amount of an agent for a time and under conditions sufficient to increase or decrease ELF5 activity.

Modulation of said activity by the administration of an agent to a mammal can be achieved by one of several techniques, including but in no way limited to introducing into said mammal a proteinaceous or non-proteinaceous molecule which:

(i) modulates expression of *ELF5*;

(ii) functions as an antagonist of ELF5;

(iii) functions as an agonist of ELF5.

Said proteinaceous molecule may be derived from natural or recombinant sources including fusion proteins or following, for example, natural product screening. Said non-proteinaceous molecule may be, for example, a nucleic acid molecule or may be derived from natural sources, such as for example natural product screening or may be chemically synthesised. The present invention contemplates chemical analogs of ELF5 capable of

acting as agonists or antagonists of ELF5. Chemical agonists may not necessarily be derived from ELF5 but may share certain conformational similarities. Alternatively, chemical agonists may be specifically designed to mimic certain physiochemical properties of ELF5. Antagonists may be any compound capable of blocking, inhibiting or otherwise preventing ELF5 from carrying out its normal biological functions. Antagonists include monoclonal antibodies specific for ELF5, or parts of ELF5, and antisense nucleic acids which prevent transcription or translation of *ELF5* genes or mRNA in mammalian cells.

Said proteinaceous or non-proteinaceous molecule may act either directly or indirectly to modulate the expression of *ELF5* or the activity of ELF5. Said molecule acts directly if it associates with *ELF5* or ELF5 to modulate the expression or activity of *ELF5* or ELF5. Said molecule acts indirectly if it associates with a molecule other than *ELF5* or ELF5 which other molecule either directly or indirectly modulates the expression or activity of *ELF5* or ELF5. Accordingly, the method of the present invention encompasses the regulation of *ELF5* or ELF5 expression or activity via the induction of a cascade of regulatory steps which lead to the regulation of *ELF5* or ELF5 expression or activity.

Another aspect of the present invention contemplates a method of modulating cellular functional activity in a mammal said method comprising administering to said mammal an effective amount of an agent for a time and under conditions sufficient to modulate the expression of a nucleotide sequence encoding *ELF5* or sufficient to modulate the activity of ELF5.

Yet another aspect of the present invention contemplates a method of modulating cellular functional activity in a mammal said method comprising administering to said mammal an effective amount of ELF5 or *ELF5*.

The ELF5, *ELF5* or agent used may also be linked to a targeting means such as a monoclonal antibody, which provides specific delivery of the ELF5, *ELF5* or agent to the target cells.

In a preferred embodiment of the present invention, the ELF5, *ELF5* or agent used in the method is linked to an antibody specific for said target cells to enable specific delivery to these cells.

Administration of the ELF5, *ELF5* or agent, in the form of a pharmaceutical composition, may be performed by any convenient means. ELF5, *ELF5* or agent of the pharmaceutical composition are contemplated to exhibit therapeutic activity when administered in an amount which depends on the particular case. The variation depends, for example, on the human or animal and the ELF5, *ELF5* or agent chosen. A broad range of doses may be applicable. Considering a patient, for example, from about 0.1 mg to about 1 mg of ELF5 or agent may be administered per kilogram of body weight per day. Dosage regimes may be adjusted to provide the optimum therapeutic response. For example, several divided doses may be administered daily, weekly, monthly or other suitable time intervals or the dose may be proportionally reduced as indicated by the exigencies of the situation. The ELF5 or agent may be administered in a convenient manner such as by the oral, intravenous (where water soluble), intranasal, intraperitoneal, intramuscular, subcutaneous, intradermal or suppository routes or implanting (e.g. using slow release molecules). With particular reference to use of ELF5 or agent, these peptides may be administered in the form of pharmaceutically acceptable nontoxic salts, such as acid addition salts or metal complexes, e.g. with zinc, iron or the like (which are considered as salts for purposes of this application). Illustrative of such acid addition salts are hydrochloride, hydrobromide, sulphate, phosphate, maleate, acetate, citrate, benzoate, succinate, malate, ascorbate, tartrate and the like. If the active ingredient is to be administered in tablet form, the tablet may contain a binder such as tragacanth, corn starch or gelatin; a disintegrating agent, such as alginic acid; and a lubricant, such as magnesium stearate.

A further aspect of the present invention relates to the use of the invention in relation to mammalian disease conditions. For example, the present invention is particularly useful, but in no way limited to, use in cancer therapy.

Accordingly, another aspect of the present invention relates to a method of treating a mammal said method comprising administering to said mammal an effective amount of an agent for a time and under conditions sufficient to modulate the expression of *ELF5* or sufficient to modulate the activity of ELF5 wherein said modulation results in modulation of cellular functional activity.

In another aspect the present invention relates to a method of treating a mammal said method comprising administering to said mammal an effective amount of ELF5 or *ELF5* for a time and under conditions sufficient to modulate cellular functional activity.

Yet another aspect of the present invention relates to the use of an agent capable of modulating the expression of *ELF5* or modulating the activity of ELF5 in the manufacture of a medicament for the modulation of cellular functional activity.

A further aspect of the present invention relates to the use of ELF5 or *ELF5* in the manufacture of a medicament for the modulation of cellular functional activity.

Still yet another aspect of the present invention relates to agents for use in modulating *ELF5* expression or ELF5 activity wherein said modulation results in modulation of cellular functional activity.

Another aspect of the present invention relates to ELF5 or *ELF5* for use in modulating cellular functional activity.

In a related aspect of the present invention, the mammal undergoing treatment may be human or an animal in need of therapeutic or prophylactic treatment.

In yet another further aspect the present invention contemplates a pharmaceutical composition comprising *ELF5*, ELF5 or an agent capable of modulating *ELF5* expression or ELF5 activity together with one or more pharmaceutically acceptable carriers and/or diluents. *ELF5*, ELF5 or said agent are referred to as the active ingredients.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions (where water soluble) and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol and liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils. The proper fluidity can be maintained, for example, by the use of a coating such as licithin, by the maintenance of the required particle size in the case of dispersion and by the use of superfactants. The preventions of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thirmerosal and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredient into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and the freeze-drying technique which yield a powder of the active ingredient plus any additional desired ingredient from previously sterile-filtered solution thereof.

When *ELF5,* ELF5 and ELF5 modulators are suitably protected they may be orally administered, for example, with an inert diluent or with an assimilable edible carrier, or

it may be enclosed in hard or soft shell gelatin capsule, or it may be compressed into tablets, or it may be incorporated directly with the food of the diet. For oral therapeutic administration, the active compound may be incorporated with excipients and used in the form of ingestible tablets, buccal tablets, troches, capsules, elixirs, suspensions, syrups, wafers, and the like. Such compositions and preparations should contain at least 1% by weight of active compound. The percentage of the compositions and preparations may, of course, be varied and may conveniently be between about 5 to about 80% of the weight of the unit. The amount of active compound in such therapeutically useful compositions in such that a suitable dosage will be obtained. Preferred compositions or preparations according to the present invention are prepared so that an oral dosage unit form contains between about 0.1 μg and 2000 mg of active compound.

The tablets, troches, pills, capsules and the like may also contain the following: A binder such as gum tragacanth, acacia, corn starch or gelatin; excipients such as dicalcium phosphate; a disintegrating agent such as corn starch, potato starch, alginic acid and the like; a lubricant such as magnesium stearate; and a sweetening agent such a sucrose, lactose or saccharin may be added or a flavouring agent such as peppermint, oil of wintergreen, or cherry flavouring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar or both. A syrup or elixir may contain the active compound, sucrose as a sweetening agent, methyl and propylparabens as preservatives, a dye and flavouring such as cherry or orange flavour. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compound may be incorporated into sustained-release preparations and formulations.

Pharmaceutically acceptable carriers and/or diluents include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption

delaying agents and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, use thereof in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

It is especially advantageous to formulate parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the mammalian subjects to be treated; each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the novel dosage unit forms of the invention are dictated by and directly dependent on (a) the unique characteristics of the active material and the particular therapeutic effect to be achieved, and (b) the limitations inherent in the art of compounding such an active material for the treatment of disease in living subjects having a diseased condition in which bodily health is impaired as herein disclosed in detail.

The principal active ingredient is compounded for convenient and effective administration in effective amounts with a suitable pharmaceutically acceptable carrier in dosage unit form as hereinbefore disclosed. A unit dosage form can, for example, contain the principal active compound in amounts ranging from 0.5 μg to about 2000 mg. Expressed in proportions, the active compound is generally present in from about 0.5 μg to about 2000 mg/ml of carrier. In the case of compositions containing supplementary active ingredients, the dosages are determined by reference to the usual dose and manner of administration of the said ingredients.

The pharmaceutical composition may also comprise genetic molecules such as a vector capable of transfecting target cells where the vector carries a nucleic acid molecule capable of modulating *ELF5* expression or ELF5 activity. The vector may, for example,

be a viral vector.

Still another aspect of the present invention is directed to antibodies to ELF5 including catalytic antibodies. Such antibodies may be monoclonal or polyclonal and may be selected from naturally occurring antibodies to ELF5 or may be specifically raised to ELF5. In the case of the latter, ELF5 may first need to be associated with a carrier molecule. The antibodies and/or recombinant ELF5 of the present invention are particularly useful as therapeutic or diagnostic agents. Alternatively, fragments of antibodies may be used such as Fab fragments. Furthermore, the present invention extends to recombinant and synthetic antibodies and to antibody hybrids. A "synthetic antibody" is considered herein to include fragments and hybrids of antibodies. The antibodies of this aspect of the present invention are particularly useful for immunotherapy and may also be used as a diagnostic tool for assessing apoptosis or monitoring the program of a therapeutic regime.

For example, ELF5 can be used to screen for naturally occurring antibodies to ELF5. These may occur, for example in some degenerative disorders.

For example, specific antibodies can be used to screen for ELF5 proteins. The latter would be important, for example, as a means for screening for levels of ELF5 in a cell extract or other biological fluid or purifying ELF5 made by recombinant means from culture supernatant fluid. Techniques for the assays contemplated herein are known in the art and include, for example, sandwich assays, ELISA and flow cytometry.

It is within the scope of this invention to include any second antibodies (monoclonal, polyclonal or fragments of antibodies) directed to the first mentioned antibodies discussed above. Both the first and second antibodies may be used in detection assays or a first antibody may be used with a commercially available anti-immunoglobulin antibody. An antibody as contemplated herein includes any antibody specific to any region of ELF5.

Both polyclonal and monoclonal antibodies are obtainable by immunization with the protein

or peptide derivatives and either type is utilizable for immunoassays. The methods of obtaining both types of sera are well known in the art. Polyclonal sera are less preferred but are relatively easily prepared by injection of a suitable laboratory animal with an effective amount of ELF5, or antigenic parts thereof, collecting serum from the animal, and isolating specific sera by any of the known immunoadsorbent techniques. Although antibodies produced by this method are utilizable in virtually any type of immunoassay, they are generally less favoured because of the potential heterogeneity of the product.

The use of monoclonal antibodies in an immunoassay is particularly preferred because of the ability to produce them in large quantities and the homogeneity of the product. The preparation of hybridoma cell lines for monoclonal antibody production derived by fusing an immortal cell line and lymphocytes sensitized against the immunogenic preparation can be done by techniques which are well known to those who are skilled in the art. (See, for example Douillard and Hoffman, Basic Facts about Hybridomas, in *Compendium of Immunology* Vol II, ed. by Schwartz, 1981; Kohler and Milstein, *Nature 256*: 495-499, 1975; *European Journal of Immunology 6*: 511-519, 1976).

In another aspect of the present invention, the molecules of the present invention are also useful as screening targets for use in applications such as the diagnosis of disorders which are regulated by ELF5. For example, screening for the levels of ELF5 protein or *ELF5* mRNA transcripts in breast or prostate tissue as an indicator of a predisposition to, or the development of, breast or prostate cancer.

Yet another aspect of the present invention contemplates a method for detecting ELF5 or *ELF5* mRNA in a biological sample from a subject said method comprising contacting said biological sample with an antibody specific for ELF5 or *ELF5* mRNA or its derivatives or homologs for a time and under conditions sufficient for an antibody-ELF5 or antibody-ELF5 mRNA complex to form, and then detecting said complex.

The presence of ELF5 may be determined in a number of ways such as by Western blotting, ELISA or flow cytometry procedures. ELF5 mRNA may be detected, for example, by *in*

situ hybridization or Northern blotting. These, of course, include both single-site and two-site or "sandwich" assays of the non-competitive types, as well as in the traditional competitive binding assays. These assays also include direct binding of a labelled antibody to a target.

Sandwich assays are among the most useful and commonly used assays and are favoured for use in the present invention. A number of variations of the sandwich assay technique exist, and all are intended to be encompassed by the present invention. Briefly, in a typical forward assay, an unlabelled antibody is immobilized on a solid substrate and the sample to be tested brought into contact with the bound molecule. After a suitable period of incubation, for a period of time sufficient to allow formation of an antibody-antigen complex, a second antibody specific to the antigen, labelled with a reporter molecule capable of producing a detectable signal is then added and incubated, allowing time sufficient for the formation of another complex of antibody-antigen-labelled antibody. Any unreacted material is washed away, and the presence of the antigen is determined by observation of a signal produced by the reporter molecule. The results may either be qualitative, by simple observation of the visible signal, or may be quantitated by comparing with a control sample containing known amounts of hapten. Variations on the forward assay include a simultaneous assay, in which both sample and labelled antibody are added simultaneously to the bound antibody. These techniques are well known to those skilled in the art, including any minor variations as will be readily apparent. In accordance with the present invention the sample is one which might contain Bim including cell extract, tissue biopsy or possibly serum, saliva, mucosal secretions, lymph, tissue fluid and respiratory fluid. The sample is, therefore, generally a biological sample comprising biological fluid but also extends to fermentation fluid and supernatant fluid such as from a cell culture.

In the typical forward sandwich assay, a first antibody having specificity for the ELF5 or antigenic parts thereof, is either covalently or passively bound to a solid surface. The solid surface is typically glass or a polymer, the most commonly used polymers being cellulose, polyacrylamide, nylon, polystyrene, polyvinyl chloride or polypropylene. The solid supports may be in the form of tubes, beads, discs of microplates, or any other surface suitable for

conducting an immunoassay. The binding processes are well-known in the art and generally consist of cross-linking covalently binding or physically adsorbing, the polymer-antibody complex is washed in preparation for the test sample. An aliquot of the sample to be tested is then added to the solid phase complex and incubated for a period of time sufficient (e.g. 2-40 minutes) and under suitable conditions (e.g. 25°C) to allow binding of any subunit present in the antibody. Following the incubation period, the antibody subunit solid phase is washed and dried and incubated with a second antibody specific for a portion of the hapten. The second antibody is linked to a reporter molecule which is used to indicate the binding of the second antibody to the hapten.

An alternative method involves immobilizing the target molecules in the biological sample and then exposing the immobilized target to specific antibody which may or may not be labelled with a reporter molecule. Depending on the amount of target and the strength of the reporter molecule signal, a bound target may be detectable by direct labelling with the antibody. Alternatively, a second labelled antibody, specific to the first antibody is exposed to the target-first antibody complex to form a target-first antibody-second antibody tertiary complex. The complex is detected by the signal emitted by the reporter molecule.

By "reporter molecule" as used in the present specification, is meant a molecule which, by its chemical nature, provides an analytically identifiable signal which allows the detection of antigen-bound antibody. Detection may be either qualitative or quantitative. The most commonly used reporter molecules in this type of assay are either enzymes, fluorophores or radionuclide containing molecules (i.e. radioisotopes) and chemiluminescent molecules.

In the case of an enzyme immunoassay, an enzyme is conjugated to the second antibody, generally by means of glutaraldehyde or periodate. As will be readily recognized, however, a wide variety of different conjugation techniques exist, which are readily available to the skilled artisan. Commonly used enzymes include horseradish peroxidase, glucose oxidase, beta-galactosidase and alkaline phosphatase, amongst others. The substrates to be used with the specific enzymes are generally chosen for the production, upon hydrolysis by the corresponding enzyme, of a detectable color change. Examples of suitable enzymes include

alkaline phosphatase and peroxidase. It is also possible to employ fluorogenic substrates, which yield a fluorescent product rather than the chromogenic substrates noted above. In all cases, the enzyme-labelled antibody is added to the first antibody hapten complex, allowed to bind, and then the excess reagent is washed away. A solution containing the appropriate substrate is then added to the complex of antibody-antigen-antibody. The substrate will react with the enzyme linked to the second antibody, giving a qualitative visual signal, which may be further quantitated, usually spectrophotometrically, to give an indication of the amount of hapten which was present in the sample. "Reporter molecule" also extends to use of cell agglutination or inhibition of agglutination such as red blood cells on latex beads, and the like.

Alternately, fluorescent compounds, such as fluorecein and rhodamine, may be chemically coupled to antibodies without altering their binding capacity. When activated by illumination with light of a particular wavelength, the fluorochrome-labelled antibody adsorbs the light energy, inducing a state to excitability in the molecule, followed by emission of the light at a characteristic color visually detectable with a light microscope. As in the EIA, the fluorescent labelled antibody is allowed to bind to the first antibody-hapten complex. After washing off the unbound reagent, the remaining tertiary complex is then exposed to the light of the appropriate wavelength the fluorescence observed indicates the presence of the hapten of interest. Immunofluorescene and EIA techniques are both very well established in the art and are particularly preferred for the present method. However, other reporter molecules, such as radioisotope, chemiluminescent or bioluminescent molecules, may also be employed.

The present invention also contemplates genetic assays such as involving PCR analysis to detect *ELF5* or its derivatives.

Further features of the present invention are more fully described in the following examples. It is to be understood, however, that this detailed description is included solely for the purposes of exemplifying the present invention. It should not be understood in any way as a restriction on the broad description of the invention as set out above.

SUMMARY OF SEQ ID Nos.

Sequence	SEQ ID NO.
nucleotide sequence of human ELF5	1
amino acid sequence of human ELF5	2
nucleotide sequence of human ELF5 short transcript	3
amino acid sequence of human ELF5 short transcript	4
nucleotide sequence of murine ELF5a	5
nucleotide sequence of murine ELF5b	6
amino acid sequence of murine ELF5	7
oligonucleotide primer	8
oligonucleotide primer	9
oligonucleotide primer	10
oligonucleotide primer	11
oligonucleotide primer	12
oligonucleotide primer	13
oligonucleotide primer	14
oligonucleotide primer	15

EXAMPLE 1

ISOLATION OF MOUSE AND HUMAN ELF5 cDNAs

The murine ELF5 cDNA was isolated from an adult mouse lung cDNA library. Amalgamation of sequence data revealed a 1437 bp sequence with a maximum open reading frame (ORF) of 759 bp, predicted to encode a 253 amino acid protein of approximately 31 kD (Figure 1a). An upstream, in-frame ston codon suggests that this ORF represents the full-length coding sequence of ELF5. Additional 91 bp of 5', and 696 bp of 3' sequences were obtained by reverse transcriptase polymerase chain reaction (PCR) and rapid amplification of cDNA ends (RACE), using day 14 mouse placental RNA. Sequence analysis revealed two discrete polyadenylation signals present in the 3' untranslated region (UTR). The first of these (at 1391 bp) appears to be an overlapping poly(A)$^+$ recognition signal, AATTAA and ATTAAAA. The second is a consensus polyadenylation signal, AATAAA, at 2181 bp. These polyadenylation signals are found close to the 3' termination of the original clone and the 3' RACE product, respectively, suggesting that these represent polyA signals for two separate mRNA products. Thus, the two predicted ELF5 cDNAs are 2224 bp and 1528 bp long. Northern blot analysis, using the ELF5 coding sequence as a probe, confirmed the presence of two predominant ELF5 transcripts in placental tissue, ELF5a and ELF5b, of approximately 2.5 kb and 1.5 kb respectively. Only ELF5a was identified using a 3' UTR fragment from between the polyadenylation signals as a probe (Figure 1b), indicating that the transcripts differ in 3' UTR sequences.

A human ELF5 cDNA fragment was isolated from a human lung cDNA library following screening with a cDNA probe containing the coding sequence of mouse ELF5. The full coding sequence of human ELF5 was then obtained by reverse transcriptase PCR and RACE using human placental RNA. Analysis revealed that the ELF5 sequence is predicted to encode a 255 residue amino acid protein.

EXAMPLE 2

COMPARISON OF HUMAN AND MOUSE ELF5 AMINO ACID SEQUENCES

The predicted amino acid sequences of human and mouse ELF5 are highly conserved, with approximately 95% identity (Figure 2a). Only a single amino acid substitution was observed within the putative ETS domain of human and mouse ELF5, and most of the other differing amino acid residues in the full-length sequences are conservative substitutions (8/13), suggesting that the two proteins are homologs (i.e. having an inferred common ancestry). Interestingly, human ELF5 does, however, contain an additional two amino acid insertion compared to mouse ELF5. In addition to the ETS domain, other features appear to be conserved between these two sequences. These include a putative 'pointed' domain (Seth *et al.*, 1992; Lautenberger *et al.*, 1992) and several consensus casein kinase II (CKII) (Pinna, 1990), protein kinase C (PKC) (Kishimoto *et al.*, 1985; Woodget *et al.*, 1986) and tyrosine kinase (Patschinsky *et al.*, 1982; Hunter, 1982; Cooper *et al.*, 1984) phosphorylation sites.

The ETS domain found within all members of the ETS family is responsible for sequence-specific DNA binding (Seth *et al.*, 1992; Lautenberger *et al.*, 1992; Wasylyk *et al.*, 1993). The putative ETS domain of human/mouse ELF5, situated at the carboxyl terminal of the protein, is similar to that of human/mouse ELF3, with amino acid identity being 67%. However, this domain is only moderately similar to that of other ETS family members, with the highest amino acid identity being 49% to human NERF, 48% to *Drosophila* ETS4 and E74A, and 46% to human ELF1 and ELK1 (Figure 2b). Sequence identity to other family members is in the range of 44-36%. However, amino acids highly conserved amongst ETS family members (Janknecht and Nordheim, 1993) are well conserved in ELF5 (23/38). Some of these highly conserved residues, such as the three tryptophan residues in the carboxyl half of the ETS domain, have been demonstrated to be structurally critical for DNA binding of other ETS family members (Wang *et al.*, 1992; Wasylyk *et al.*, 1992).

Based on ETS domain similarities, a recent phylogenetic analysis (Graves and Petersen, 1998) has proposed the grouping of ETS factors into subfamilies, one of which is the ELF (E-74-like-factor) subfamily. The ELF subfamily includes *Drosophila* E74A, human ELF1 and NERF. A phylogenetic tree was generated including ELF5 and recently isolated ELF3, by maximum likelihood analysis of the ETS domain (Figure 2c). It shows that the human and mouse ELF5 sequences group most closely with the human and mouse ELF3 sequences, and that both ELF3 and ELF5 are most closely related to *Drosophila* ETS4, E74A and human ELF1 and NERF within the ETS family. Thus, *Drosophila* ETS4, and human/mouse ELF3 and ELF5 may also fall into the ELF subfamily of ETS factors.

The phylogency in Figure 2c shows the unrooted relationship among 28 ETS domains.

EXAMPLE 3

HUMAN CHROMOSOMAL MAPPING OF ELF5

Human chromosomal localization of ELF5 was performed by PCR, using gene specific primers and the Genebridge 4 Radiation Hybrid DNA panel (UK HGMP Resource Centre). With these primers, a single product of the expected size (725 bp) was amplified from total human DNA. The PCR reactions were then performed separately for each of the individual hybrids. The amplification results from the 93 hybrids were submitted to the Radiation Hybrid Mapping server at Whitehead Institute/MIT Center for Genome Research for analysis. The result demonstrated that ELF5 is localised to chromosome 1. The marker most tightly linked to ELF5 was NIB1364 (D1S3023) at a distance of OcR (lod score > 3.0), and this marker is located in the region of 1p36.31 (Figure 3). This chromosomal region frequently undergoes loss of heterozygosity (LOH) in several types of carcinoma (Genuardi *et al.*, 1989; Praml *et al.*, 1995; Takayama *et al.*, 1993; Thrash-Bingham *et al.*, 1996; Wada *et al.*, 1998).

EXAMPLE 4

EXPRESSION PATTERN OF ELF5 IN MOUSE TISSUES

Poly(A)$^+$ mRNA material derived from various mouse tissues were analysed by Northern blot hybridization using the murine ELF5 cDNA as a probe. A GAPDH probe was then used to control for RNA loading.

Analysis of ELF5 expression in adult mouse tissues revealed that ELF5 has a restricted expression pattern. Expression of two ELF5 transcripts, ELF5-a (2.5 kb) and ELF5-b (1.5 kb), were observed in lung (Lu), kidney (Ki), stomach (St), ovary (Ov), tongue (To), bladder (Bl), and day 2 pregnant (2 Ma) and day 10 pregnant (10 Ma) mammary glands, but no expression was observed in liver (Li), heart (He), small intestine (Sm), spleen (Sp), thymus (Th), pancreas (Pa), skeletal muscle (Sk), colon (Co) or fat (2 Fa and 10 Fa) (Figure 4a). Fat from day 2 (2 Fa) and day 10 (10 Fa) pregnant mice was used as a control for mammary expression, since the mammary gland contains much fat tissue. A single transcript was observed in brain (arrow - approximately 2.1 kb), but of a different size to either of the two ELF5 transcripts in other organs.

The expression of ELF5 was examined in the neonatal mouse (Figure 4b) and during embryogenesis on days 19, 17 and 16 (Figure 4c), and observed a similar expression pattern compared to that of the adult. However, at day 16 stage of embryogenesis low levels of ELF5 expression were detected in brain (regular sized transcripts) and small intestine, in addition to the expression pattern observed in the adult.

Placental expression of ELF5 displayed an interesting pattern during stages of embryogenesis (Figure 4d). Both transcripts were increasingly expressed from day 9.5 to day 13 before an overall decrease observed from day 14 to day 19, although some expression was observed at day 17.

The two predominant ELF5 mRNA transcripts were observed in variable ratios in

different tissues, suggesting that polyadenylation sites may be utilized differentially, or the two transcripts are subject to differential degradation. ELF5-a was expressed more strongly in neonatal and embryonic lung and kidney (Figures 4b and c), and adult ovary (Figure 4a), compared to ELF5-b. Conversely, ELF5-b was stronger in adult tongue (Figure 4a), and in all developmental stages of stomach (Figures 4a, b and c), compared to ELF5-a. In some RNA samples a further large (>10 kb) transcript was variably observed.

EXAMPLE 5

EXPRESSION PATTERN OF ELF5 IN HUMAN TISSUES AND CANCER CELL LINES

Expression of ELF5 in adult human organs was also analysed by Northern blot of poly(A)$^+$ mRNA probed with the human ELF5 cDNA (Figure 5a). A single transcript of approximately 2.5 kb was strongly expressed in kidney (Ki) and prostate (Pr). However, much longer exposures of blots demonstrated just detectable expression of ELF5 in placenta (Pl) and lung (Lu). Further, ELF5 was cloned from human lung and placenta cDNA libraries, confirming that it is expressed in these tissues, albeit probably at very low levels.

ELF5 expression in human cancers was examined. A panel of cancer cell lines, including carcinomas of the ovary (CaOv-3), breast (BT-549, ZR-75-1, T47D), kidney (786-0), liver (SK-HEP-1), lung (A549), amnion (WISH), prostate (DU145, PC3) and endometrium (HEC-1), and melanoma (MEL28), T-cell leukemia (Jurkat) and erythroid leukemia (K562), were analysed for ELF5 expression by RNAse protection assay (Figure 5b). A primary fibroblast cell line (CCL32SK) was also included as a sample of non-transformed cells. Of all these cell lines only T47D, a progesterone sensitive ductal breast carcinoma, was observed to express ELF5.

To evaluate the possibility that lack of ELF5 expression in carcinoma was due to

genomic alterations, a panel of breast and lung carcinoma derived cell lines were analysed by Southern blot (Figure 5c). ELF5 gene dosage was compared to that present in DNA from normal human blood (based on the 6.5 kb *Bgl*II fragment) and controlled by hybridization with a β-*actin* cDNA probe. These results are summarized in the lower panel, where '2' represents a normal allele complement. No evidence was found for allelic loss or gene rearrangement in the two breast carcinoma cell lines that did not express ELF5 (BT-549 - lane 2, ZR-75-1 - lane 3). However, of nine lung carcinoma cell lines, evidence for loss of an ELF5 allele was observed in two (NCI-H358 - lane 8, NCI-H441 - lane 11). Hybridization with an ELF3 cDNA probe, which is localised to the long arm of chromosome 1 (Tymms *et al.*, 1997), helped to confirm the specific loss of ELF5 alleles. Two other lung carcinoma lines (SK-LU-1 - lane 10, NCI-H661 - lane 13) displayed hybridization with multiple fragments (shaded arrows) in addition to those observed in normal DNA (solid arrows), possibly indicating that at least one ELF5 allele has been rearranged in these lines. Confirmation of rearrangement, rather than restriction fragment length polymorphism (RFLP), was made by additional restriction digests. Some cell lines appeared to have amplification or additional copies of the ELF5 gene. One of these, T47D (lane 4), was the only cell line demonstrated to express ELF5, and another, SK-LU-1 (lane 10), appeared to have rearranged alleles.

EXAMPLE 6

SEQUENCE-SPECIFIC BINDING OF ELF5 TO DNA SEQUENCES CONTAINING CONSENSUS ETS SITES

Although ELF5 displays similarity to the consensus ETS domain, characterising it as an ETS family member, this sequence is still quite divergent from most other ETS family members. The hallmark of ETS factors to bind DNA sites containing a GGAA-core in a sequence-specific manner is however shared by ELF5, demonstrating an additional functional similarity to the ETS family. A recombinant ELF5 HIS-tag protein of approximately 29 kD, expressed in *E. coli* and purified by metal-affinity chromatography (Figure 6a, lane 4), displayed strong binding to consensus ETS binding sites, as analysed

by electrophoretic mobility shift assay (EMSA) (Figure 6b). ELF5 bound the E74 oligonucleotide (containing a GGAA-core) (lane 1), but not to the E74ml oligonucleotide (which had been mutated to an AGAA-core) (lane 2). The first G-residue of the core has been demonstrated to be a physical point of DNA contact for ETS1, and consequently essential for DNA binding (Fisher *et al.*, 1991; Nye *et al.*, 1992). Thus, ELF5 displays sequence specific binding to a consensus ETS binding site, binding that is disrupted by a mutation known to similarly affect other ETS family members. These results were confirmed through competition analysis. The ELF5-E74 complex (lane 3) was efficiently competed by the addition of a 100-fold excess of unlabeled E74 (lane 4), but not by E74ml (lane 5).

ELF5 also displayed sequence specific binding to different consensus ETS binding sequences, and did so with differential affinity (Figure 6b). Competition of the ELF5-E74 complex (lane 3) was achieved by consensus ETS sites from the GM-CSF promoter (lane 6), *erb*-B2 promoter (lane 7) and moloney sarcoma virus (MSV) long terminal repeat (LTR) (lane 8). The relative ability of ELF5 to bind these sequences occurred in the order: E74 > erbB2 > MSV > GM-CSF. ELF5 did not appear to be competed at all by an oligonucleotide containing a consensus AP1 binding site (lane 9). ETS1 binding to E74 was used as a positive control (lane 10).

EXAMPLE 7

MOUSE ELF5 ACTS AS A TRANSCRIPTIONAL ACTIVATOR

In addition to DNA binding, another characteristic of most ETS factors is their ability to transactivate from binding sites in promoters and enhancers.

A reporter construct, containing the chloramphenicol acetyl-transferase (CAT) driven by a minimal TK promoter and multiple ETS/AP1 binding sites (from the polyomavirus enhancer), was co-transfected into COS cells together with an ELF5 expression construct (Figure 7). Analysis of CAT activities revealed that ELF5 expression resulted in an

average five-fold transactivation of the reporter. Further, this transactivation was inhibited by addition of an anti-sense ELF5 mRNA expression vector, indicating that ELF5 transactivation was due specifically to the product translated from the sense construct.

EXAMPLE 8

ISOLATION AND CHARACTERIZATION OF FULL-LENGTH MURINE ELF5 CDNA

The murine *Elf5* cDNA was isolated from an adult lung cDNA library in Lambda ZAPII (Stratagene) following screening with a cDNA probe containing the ETS domain region of human *ELF3*. Additional 5' sequence and 3' sequence were obtained by RT-PCR using a Marathon cDNA synthesis Kit (Clontech) and RACE (Rapid Amplification using day 14 murine of cDNA Ends) placental Poly(A)$^-$ RNA. The murine *Elf5*-specific PCR products were cloned into pGEM-T vector (Promega Corp., Madison, WI, USA). All cDNA sequences were confirmed by sequencing both strands at least once. 5'RACE gene-specific primer 1:5'GCCAGTCTTG-GTCTCTTCAGCATC-3' (SEQ ID NO:8); 5'-RACE nested-gene-specific primer 2: 5'-AGGAGATGCAGTTGGCATCAAGCT-3' (SEQ ID NO:9); 3'-RACE gene-specific primer 1: 5'-AGCCAGTGTTATGGGTGCTG-3' (SEQ ID NO:10); 3'-RACE nested-gene-specific primer 2: 5'-ACAGTCACTTGATCCACGGCCAATCC-3' (SEQ ID NO:11).

EXAMPLE 9

ISOLATION OF HUMAN ELF5 CODING SEQUENCE

A human *ELF5* cDNa fragment was isolated from a human lung cDNA library (GIBCO BRL) following screening with a cDNA probe containing the coding sequence of mouse *Elf5*. The coding sequence was then obtained by RT-PCR using a Marathon cDNA synthesis Kit (Clontech) and RACE (Rapid Amplification of cDNA Ends) using human placental Poly(A)$^+$ RNA. The human *ELF5*-specific PCR products were cloned into pGEM-T vector (Promega Corp., Madison, WI, USA). All cDNA sequences were

confirmed by sequencing both strands at least once.

EXAMPLE 10

STS CONTENT MAPPING

The following sequence specific primers for human ELF5 were used for PCR. Forward primer: 5'-GGGTGGCAGGAAGACAAGCTATGA-3 (SEQ ID NO:12); Reverse primer: 5'-CCAATTAAGTCCCAGCTTGATGGC-3 (SEQ ID NO:13). The PCR reactions were performed in Opti-Primer™ 10 x buffer #3 (100 mM Tris-HCl pH 8.3, 35 mM $MgCl_2$, 250 mM Kcl) with 1 μl of Master Mix 50 x buffer (20 mM Tris-HCl pH 8.0, 250 nM EDTA) (Opti-Primer™ PCR Optimization Kit, Stratagene), 50 ng of template DNA, 0.2 μg of each primer, 1 μl of 10 mM dNTPs and 0.25 U of Taq DNA polymerase in a total volume of 50 μl. PCR parameters were an initial denaturation step at 94°C for 1 min, followed by 30 cycles of 94°C (1 min), 65°C (1 min), 72°C (1 min). For Genebridge 4 Radiation Hybrid DNA panel (UK GHMP Resource Centre), PCR reactions were performed separately for each of the individual hybrids. The PCR results from the 93 hybrids were submitted to the Radiation Hybrid Mapping server at Whitehead Institute/MIT Center for Genome Research (http://www.genome.wi.mit.edu/cgibin/contig/rhmapper.pl). The STS content mapping experiment was performed in duplicate and included PCR reactions with no DNA, total human DNA and total hamster DNA as controls.

EXAMPLE 11

SOUTHERN AND NORTHERN BLOT ANALYSIS

Northern analysis of *ELF5* expression in human adult organs was performed with commercially available blots containing 2 μg of Poly(A)$^+$ RNA (Clontech). For other Northern blots POLY(A)$^+$ mRNA was isolated by a modification of Gonda *et al.* (1992). Genomic DNA was isolated by standard techniques (Sambrook *et al*, 1997). Random-primed probes using a 898 bp human *ELF5* cDNA fragment and a 940 bp *StyI* mouse *Elf5* cDNA fragment were generated and Southern/Northern hybridizations performed using standard procedures. Blots were re-probed with glyceraldehyde-3-phosphate

dehydrogenase (GAPDH) or β-*actin* cDNAs to verify RNA/DNA loading.

EXAMPLE 12
RNASE PROTECTION ANALYSIS

ELF5 mRNA abundance in total RNA from human cell lines was determined as described previously (Tymms, 1995). Anti-sense RNA probes for human *ELF5* and GAPDH transcribed from linearized plasmid vectors generated full-length probes of 388 bp and 216 bp, respectively. The protected products generated by hybridization and RNAse digestion are 298 bp for *ELF5* and 150 bp for GAPDH.

EXAMPLE 13
CELL LINES AND CULTURE

Monkey COS7 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum, 100 U/ml penicillin and 100 μg/ml streptomycin, and maintained in a humidified incubator at 5% CO_2 and 37°C.

EXAMPLE 14
PLASMIDS

pHis6-Elf5 expression vector was made as follows: The murine *Elf5* cDNA was amplified using PCR oligonucleotide primers (5'-CGGGATCCTTGGACTCCGTAACCCATAGC-3' (SEQ ID NO:14) and 5'-GCAGATCTCAGAGTTTCTCTTCCTGCC-3' (SEQ ID NO:15)) containing a *Bam*HI restriction site followed by 21 nucleotides of the murine *Elf5* coding sequence and a *Bgl*II restriction site followed by 19 nucleotides complementary to the last 20 nucleotides of the *Elf5* coding sequence. The PCR fragment was cloned into the pGEM-T vector (Promega Corp., Madison, WI, USA), the *Bam*HI-*Sac*I restriction fragment with the *Elf5* coding sequence was then cloned into the *Bam*HI-*Sac*I sites of the pQE30 (Qiagen, Inc. Chatsworth, CA, USA) bacterial expression vector resulting in a N-terminal fusion of *Elf5* protein to six histidine residues (His-Tag).

The *Elf5* mammalian expression construct (pBOSElf5s) contains the full mouse *Elf5* cDNA blunt cloned into the T4 polymerase blunted *Xba*I site of pEFBOS (Mizushima and Natata, 1990). Expression from pEFBOS is driven by the elongation factor-1 promoer. The *Elf5* anti-sense expression construct is similar, but with reverse orientation of the *Elf5* polyomavirus enhancer oligonucleotides into the *Bam*HI site of pBLCAT2.

EXAMPLE 15
ELECTROPHORETIC MOBILITY SHIFT ASSAYS (EMSA)

Purified recombinant *Elf5* and Ets1 proteins were produced as 6XHis-tag fusions in *E. coli* using the QIAexpress expression system (Qiagen). Overnight cultures were diluted 1/10 in LB broth and grown for 1 h at 37°C. Expression of recombinant proteins were induced by addition of 0.1 mM IPTG and culture of cells for 2 h. Cells were harvested and sonicated in lysis buffer (6 M guanidine, 20 mM Tris-HCl, 50 mM NaCl, pH 8.0), and cell debris removed by centrifugation. One ml of metal His-affinity resin was incubated with supernatants for 30 min, collected, washed in wash buffer (8 M urea, 20 mM Tris-HCl, 50 mM NaCl, pH 8.0), and resuspended in renaturation buffer (20 mM Tris-HCl, 50 mM NaCl, 3 mM dithiothreitol (DTT), pH 8.0). Proteins were eluted from the beads in renaturation buffer supplemented with 100 mM imidazole. Purification and integrity of recombinant proteins were confirmed by denaturing SDS-polyacrylamide gel electrophoresis (PAGE).

DNA binding experiments with recombinant proteins were performed using EMSA, as previously described (Thomas *et al*, 1995, 1997). Briefly, purified double stranded oligonucleotides were labeled with γ-^{32}P dATP and T4 polynucleotide kinase. Oligonucleotide probe (1 ng) was incubated for 10 min with approximately 20 ng purified *Elf5/Ets1* protein in DNA binding buffer (1 mM EDTA, 10 mM Tris-Hcl pH 8.0, 50 mM NaCl, 3 mM DTT, 1 mg/ml BSA, 500 ng/ml poly-d(I-C)d(I-C), 500 ng/ml poly dI-dC, 200 ng/ml sheared salmon sperm DNA), ± 100 ng unlabelled competitor oligonucleotides, in 10 μl final volume. Assays were run through non-denaturing, 7% acrylamide (29 acrylamide:1 bis-acrylamide), 0.5 x TBE gels at 4°C.

EXAMPLE 16

LACK OF ELF5 EXPRESSION IN HUMAN PRIMARY BREAST CARCINOMAS

A panel of human primary breast carcinoma samples were analysed for ELF5 expression by *in situ* hybridization (Hogan *et al.*, 1994). Section from parafin-embedded samples were hybridized with ELF5 ^{33}P-labelled antisense RNA, and signals were detected with a photosensitive emulsion. Serial sections were also stained with eosin and haematoxylin.

A preliminary examination of ELF5 expression shows that ELF5 is not detectable in 20 out of 20 human breast carcinomas studied, whereas it is strongly expressed in adjacent normal epithelium and in epithelial cells from normal subjects (Figure 8).

BIBLIOGRAPHY

Adachi J and Hasegawa M., *MOLPHY Version 2.3: Programs for molecular phylogenetics based on maximum likelihood.* Computer Science Monographs. The Institute of Statistical Mathematics, Tokyo

Ben-David Y., Giddens EB., Letwin K and Bernstein A., *Genes. Dev.*, **5**:908-918 (1991)

Bhat NK, Komschlies KL, Fujiwara S, Fisher RJ, Mathieson BJ, Gregorio TA, Young HA, Kasik JW, Ozato K and Papas TS., *J. Immunol.*, **142**:672-678 (1989)

Bhat NK, Thompson CB, Lindsten T, June CH, Fujiwara S, Koizumi S, Fisher RJ and Papas TS., *Proc. Natl. Acad. Sci. USA.*, **87**:3723-3727 (1990)

Buttice G and Kurkinen M., *J. Biol. Chem.*, **268**:7196-7204 (1993)

Buttice G, Duterque-Coquillaud M, Basuyaux JP, Carrere S, Kurkinen M and Stehelin D., *Oncogene.*, **13**:2297-2306 (1996)

Chang CH, Scott GK, Kuo WL, Xiong X, Suzdaltseva Y, Park JW, Sayre P, Erny K, Collins C, Gray JW and Benz CC., *Oncogene.*, **14**:1617-1622 (1997)

Choi SG, Yi Y, Kim YS, Kato M, Chang J, Chung HW, Hahm KB, Yang HK, Rhee HH, Bang YJ and Kim SJ., *J. Biol. Chem.*, **273**:110-117 (1998)

Cooper JA, Esch FS, Taylor SS and Hunter T., *J. Biol. Chem.*, **259**:7835-7841 (1984)

Delannoy-Courdent A, Fauquette W, Dong-Le Bourhis XF, Boilly B, Vandenbunder B and Desbiens X., *Int. J. Dev. Biol.*, **40**:1097-1108 (1996)

Fisher RJ, Mavrothalassitis G, Kondoh A and Papas TS., *Oncogene.*, **6**:2249-2254 (1991)

Genuardi M, Tsihira H, Anderson DE and Saunders GF., *Am. J. Hum. Genet.*, **45**:73-82 (1989)

Golub TR, Barker GF, Bohlander SK, Hiebert SW, Ward DC, Bray-Ward P, Morgan E, Raimondi SC, Rowley JD and Gilliland DG., *Proc. Natl, Acad. Sci. USA.*, **92**:4917-4921 (1995)

Golub TR, Barker GF, Lovett M and Gilliland DG., *Cell.*, **77**:307-316 (1994)

Gonda TJ, Sheiness DK and Bishop JM., *Mol. Cell Biol.*, **2**:617-624 (1992)

Graves BL and Petersen JM., *Adv. Cancer Res.*, in press (1998)

Gutman A and Wasylyk B., *Trends Genet.*, 7:49-54 (1991)

Hart AH, Corrick CM, Tymms MJ, Hertzog PJ and Kola I., *Oncogene.*, **10**:1423-1430 (1995)

Hogan B, Beddington R, Costantini F and Lacy E., *A Laboratory Manual. 2nd edn. Cold Spring Harbor Laboratory Press.*, **p**.344-351 (1994)

Hunter T., *J. Biol. Chem.*, **257**:4843-4848 (1982)

Ida K, Kobayashi S, Taki T, Hanada R, Bessho F, Yamamori S, Sugimoto T, Ohki M and Hayashi Y., *Int. J. Cancer.*, **63**:500-504 (1995)

Janknecht R and Nordheim A., *Biochim. Biophys. Acta.*, **1155**:346-356 (1993)

Jones DT, Taylor WR and Thornton JM., *Nature.*, **358**:86-89 (1992)

Karim FD, Urness LD, Thummel CS, Klemsz MJ, McKercher SR, Celada A, Van Beveren C, Maki RA, Gunther CV, Nye JA and Graves BJ., *Genes Dev.*, **4**:1451-1453 (1990)

Kishimoto A, Nishiyama K, Nakanishi H, Uratsuji Y, Nomura H, Takeyama Y and Nishizuka Y., *J. Biol. Chem.*, **260**:12492-12499 (1985)

Kodandapani R, Pio F, Ni CZ, Piccialli G, Klemsz M, McKercher S, Maki RA and Ely KR., *Nature.*, **380**:456-460 (1996)

Kola I, Brookes S, Green AR, Garber R, Tymms M, Papas TS and Seth A., *Proc. Natl. Acad. Sci. USA.*, **90**:7588-7592 (1993)

Lautenberger JA, Burdett LA, Gunnell MA, Qi S, Watson DK, O'Brien SJ and Papas TS., *Oncogene.*, 7:1713-1719 (1992)

Macleod K, Leprince D and Stehelin D., *Trends Biochem. Sci.*, **17**:251-256 (1992)

Mizushima S and Nagata S., *Nucleic Acids Res.*, **18**:5322 (1990)

McKercher SR, Torbett BE, Anderson KL, Henkel GW, Vestal DJ, Baribault H, Klemsz M, Feeney AJ, Wu GE, Paige CJ and Maki RA., *EMBO J.*, **15**:5647-5658 (1996)

Moreau-Gachelin F, Wendling F, Molina T, Denis N, Titeux M, Grimber G, Briand P, Vainchenker W and Tavitian A., *Mol. Cell Biol.*, **16**:2453-2463 (1996)

Muthusamy N, Barton K and Leiden JM., *Nature.*, **377**:639-642 (1995)

Nunn MF, Seeburg PH, Moscovici C and Duesberg PH., *Nature.*, **306**:391-395 (1983)

Nye JA, Petersen JM, Gunther CV, Jonsen MD and Graves BJ., *Genes Dev.*, **6**:975-990 (1992)

O'Neill EM, Rebay I, Tjian R and Rubin GM., *Cell.*, **78**:137-147 (1994)

Oettgen P, Alani RM, Barcinski MA, Brown L, Akbarali Y, Boltax J, Kunsch C, Munger K and Libermann TA., *Mol. Cell Biol.*, **17**:4419-4433 (1997)

Patschinsky T, Hunter T, Esch FS, Cooper JA and Sefton BM., *Proc. Natl. Acad. Sci. USA.*, **79**:973-977 (1982)

Peeters P, Raynaud SD, Cools J, Wlodarska I, Grosgeorge J, Philip P, Monpoux F, Van Rompaey L, Baens M, Van den Berghe H and Marynen P., *Blood.*, **90**:2535-2540 (1997)

Peter M, Couturier J, Pacquement H, Michon J, Thomas G, Magdelenat H and Delattre O., *Oncogene.*, **14**:1159-1164 (1997)

Pinna LA., *Biochim. Biophys. Acta.*, **1054**:267-284 (1990)

Praml C, Finke LH, Herfarth C, Schlag P, Schwab M and Amler L., *Oncogene.*, **11**:1357-1362 (1995)

Sambrook J, Fritsch ER and Maniatis T., *Molecular Cloning: A Laboratory Manual 2nd edn. Cold Spring Harbor Laboratory Press.*, (1997)

Scott EW, Simon MC, Anastasi J and Singh H., *Science.*, **265**:1573-1577 (1994a)

Scott GK, Daniel JC, Xiong X, Maki RA, Kabat D and Benz CC., *J. Biol. Chem.*, **269**:19848-19858 (1994b)

Seth A and Papas TS., *Oncogene.*, **5**:1761-1767 (1990)

Seth A, Ascione R, Fisher RJ, Mavrothalassitis GJ, Bhat NK and Papas TS., **3**:327-334 (1992)

Seth A, Watson DK, Blair DG and Papas TS., *Proc. Natl. Acad. Sci. USA.*, **86**:7833-7837 (1989)

Sorensen PH, Lessnick SL, Lopez-Terrada D, Liu XF, Triche TJ and Denny CT., *Nat. Genet.*, **6**:146-151 (1994)

Sumarsono SH, Wilson TJ, Tymms MJ, Venter DJ, Corrick CM, Kola R, Lahoud MH, Papas TS, Seth A and Kola I., *Nature.*, **379**:534-537 (1996)

Sun W, Graves BJ and Speck NA., *J. Virol.*, **69**:4941-4949 (1995)

Takayama H, Suzuki T, Mugishima H, Fujisawa T, Ookuni M, Schwab M, Gehring M, Nakamura Y, Sugimura T, Terada M and Yokota J., *Oncogene*, 7:1185-1189 (1992)

Thomas RS, Tymms MJ, Mckinlay LH, Shannon MF, Seth A and Kola I., *Oncogene*, **14**:2845-2855 (1997)

Thomas RS, Tymms MJ, Seth A, Shannon MF and Kola I., *Oncogene*, **11**:2135-2143 (1995)

Thrash-Bingham CA, Salazar H, Greenberg RE and Tartof KD., *Genes Chromosomes Cancer*, **16**:64-67 (1996)

Tymms MJ, *Methods Mol. Biol.*, **37**:31-46 (1995)

Tymms MJ, Ng AY, Thomas RS, Schutte BC, Zhou J, Eyre HJ, Sutherland GR, Seth A, Rosenberg M, Papas T, Debouck C and Kola I, *Oncogene*, **15**:2449-2462 (1997)

Vandenbunder B, Wenert N, Queva C, Desbiens X and Stehelin D, *Folia. Biol. (Praha.)*, **40**:301-313 (1994)

Wada M, Yokota J, Mizoguchi H, Sugimura T and Terada M, *Cancer Res.*, **48**:2988-2992 (1988)

Wang CY, Petryniak B, Ho IC, Thompson CB and Leiden JM, *J. Exp. Med.*, **175**:1391-1399 (1992)

Wasylyk B, Hahn SL and Giovane A, *Eur. J. Biochem.*, **211**:-7-18 (1993)

Wasylyk C, Gutman A, Nicholson R, and Wasylyk B, *EMBO. J.*, **10**:1127-1134 (1991)

Wasylyk C, Kerckaert JP and Wasylyk B, *Genes. Dev.*, **6**:965-974 (1992)

Watson DK, McWilliams MJ, Lapis P, Lautenberger JA, Schweinfest CW and Papas TS, *Proc. Natl. Acad. Sci. USA*, **85**:7862-7866 (1988)

Werner MH, Clore M, Fisher CL, Fisher RJ, Trinh L, Shiloach J and Gronenborn AM, *Cell*, **83**:761-771 (1995)

Wernert N, Raes MB, Lassalle P, Dehouck MP, Gosselin B, Vandenbunder B and Stehelin D. *Am. J. Pathol.*, **140**:119-127 (1992)

Woodget JR, Gould KL and Hunter T, *Eur. J. Biochem.*, **161**:177-184 (1986)

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1676 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 110..875

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
CAAGGCTACA GGTGTCTTTA TTTCCACTGC ACGCTGGTGC TGGGAGCGCC TGCCTTCTCT      60

TGCCTTGAAA GCCTCCTCTT TGGACCTAGC CACCGCTGCC CTCACGGTA ATG TTG         115
                                                      Met Leu
                                                        1

GAC TCG GTG ACA CAC AGC ACC TTC CTG CCT AAT GCA TCC TTC TGC GAT       163
Asp Ser Val Thr His Ser Thr Phe Leu Pro Asn Ala Ser Phe Cys Asp
      -   5                  10                  15

CCC CTG ATG TCG TGG ACT GAT CTG TTC AGC AAT GAA GAG TAC TAC CCT       211
Pro Leu Met Ser Trp Thr Asp Leu Phe Ser Asn Glu Glu Tyr Tyr Pro
     20                  25                  30

GCC TTT GAG CAT CAG ACA GCC TGT GAC TCA TAC TGG ACA TCA GTC CAC       259
Ala Phe Glu His Gln Thr Ala Cys Asp Ser Tyr Trp Thr Ser Val His
 35                  40                  45                  50

CCT GAA TAC TGG ACT AAG CGC CAT GTG TGG GAG TGG CTC CAG TTC TGC       307
Pro Glu Tyr Trp Thr Lys Arg His Val Trp Glu Trp Leu Gln Phe Cys
                 55                  60                  65

TGC GAC CAG TAC AAG TTG GAC ACC AAT TGC ATC TCC TTC TGC AAC TTC       355
Cys Asp Gln Tyr Lys Leu Asp Thr Asn Cys Ile Ser Phe Cys Asn Phe
             70                  75                  80

AAC ATC AGT GGC CTG CAG CTG TGC AGC ATG ACA CAG GAG GAG TTT GTC       403
Asn Ile Ser Gly Leu Gln Leu Cys Ser Met Thr Gln Glu Glu Phe Val
         85                  90                  95

GAG GCA GCT GGC TTC TGC GGC GAG TAC CTG TAC TTC ATC CTC CAG AAC       451
Glu Ala Ala Gly Phe Cys Gly Glu Tyr Leu Tyr Phe Ile Leu Gln Asn
    100                 105                 110

ATC CGC ACA CAA GGT TAC TCC TTT TTT AAT GAC GCT GAA GAA AGC AAG       499
Ile Arg Thr Gln Gly Tyr Ser Phe Phe Asn Asp Ala Glu Glu Ser Lys
115                 120                 125                 130

GCC ACC ATC AAA GAC TAT GCT GAT TCC AAC TGC TTG AAA ACA AGT GGC       547
Ala Thr Ile Lys Asp Tyr Ala Asp Ser Asn Cys Leu Lys Thr Ser Gly
                135                 140                 145

ATC AAA AGT CAA GAC TGT CAC AGT CAT AGT AGA ACA AGC CTC CAA AGT       595
Ile Lys Ser Gln Asp Cys His Ser His Ser Arg Thr Ser Leu Gln Ser
            150                 155                 160

TCT CAT CTA TGG GAA TTT GTA CGA GAC CTG CTT CTA TCT CCT GAA GAA       643
Ser His Leu Trp Glu Phe Val Arg Asp Leu Leu Leu Ser Pro Glu Glu
        165                 170                 175
```

```
AAC TGT GGC ATT CTG GAA TGG GAA GAT AGG GAA CAA GGA ATT TTT CGG        691
Asn Cys Gly Ile Leu Glu Trp Glu Asp Arg Glu Gln Gly Ile Phe Arg
    180                 185                 190

GTG GTT AAA TCG GAA GCC CTG GCA AAG ATG TGG GGA CAA AGG AAG AAA        739
Val Val Lys Ser Glu Ala Leu Ala Lys Met Trp Gly Gln Arg Lys Lys
195                 200                 205                 210

AAT GAC AGA ATG ACA TAT GAA AAG TTG AGC AGA GCC CTG AGA TAC TAC        787
Asn Asp Arg Met Thr Tyr Glu Lys Leu Ser Arg Ala Leu Arg Tyr Tyr
                215                 220                 225

TAT AAA ACA GGA ATT TTG GAG CGG GTT GAC CGA AGG TTA GTG TAC AAA        835
Tyr Lys Thr Gly Ile Leu Glu Arg Val Asp Arg Arg Leu Val Tyr Lys
            230                 235                 240

TTT GGA AAA AAT GCA CAC GGG TGG CAG GAA GAC AAG CTA T GATCTGCTCC       885
Phe Gly Lys Asn Ala His Gly Trp Gln Glu Asp Lys Leu
        245                 250                 255

AGGCATCAAG CTCATTTTAT GGATTTCTGT CTTTTAAAAC AATCAGATTG CAATAGACAT      945

TCGAAAGGCT TCATTTTCTT CTCTTTTTTT TTAACCTGCA AACATGCTGA TAAAATTTCT     1005

CCACATCTCA GCTTACATTT GGATTCAGAG TTGTTGTCTA CGGAGGGTGA GAGCAGAAAC     1065

TCTTAAGAAA TCCTTTCTTC TCCCTAAGGG GATGAGGGGA TGATCTTTTG TGGTGTCTTG     1125

ATCAAACTTT ATTTTCCTAG AGTTGTGGAA TGACAACAGC CCATGCCATT GATGCTGATC     1185

AGAGAAAAAC TATTCAATTC TGCCATTAGA GACACATCCA ATGCTCCCAT CCCAAAGGTT     1245

CAAAAGTTTT CAAATAACTG TGGCAGCTCA CCAAAGGTGG GGGAAAGCAT GATTAGTTTG     1305

CAGGTTATGG TAGGAGAGGG TGAGATATAA GACATACATA CTTTAGATTT TAAATTATTA     1365

AAGTCAAAAA TCCATAGAAA AGTATCCCTT TTTTTTTTTT TGAGACGGGT TCTCACTATG     1425

TTGCCCAGGG CTGGTCTTGA ACTCCTATGC TCAAGTGATC CTCCCACCTC GGCCTCCCAA     1485

AGTACTGTGA TTACAAGCGT GAGCCACGGC ACCTGGGCAG AAAAGTATCT TAATTAATGA     1545

AAGAGCTAAG CCATCAAGCT GGGACTTAAT TGGATTTAAC ATAGGTTCAC AGAAAGTTTC     1605

CTAACCAGAG CATCTTTTTG ACCACTCAGC AAAACTTCCA CAGACATCCT TCTGGACTTA     1665

AACCGGAATT C                                                          1676
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 255 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Leu Asp Ser Val Thr His Ser Thr Phe Leu Pro Asn Ala Ser Phe
 1               5                  10                  15

Cys Asp Pro Leu Met Ser Trp Thr Asp Leu Phe Ser Asn Glu Glu Tyr
            20                  25                  30
```

```
Tyr Pro Ala Phe Glu His Gln Thr Ala Cys Asp Ser Tyr Trp Thr Ser
          35                  40                  45

Val His Pro Glu Tyr Trp Thr Lys Arg His Val Trp Glu Trp Leu Gln
      50                  55                  60

Phe Cys Cys Asp Gln Tyr Lys Leu Asp Thr Asn Cys Ile Ser Phe Cys
 65                  70                  75                  80

Asn Phe Asn Ile Ser Gly Leu Gln Leu Cys Ser Met Thr Gln Glu Glu
                 85                  90                  95

Phe Val Glu Ala Ala Gly Phe Cys Gly Glu Tyr Leu Tyr Phe Ile Leu
            100                 105                 110

Gln Asn Ile Arg Thr Gln Gly Tyr Ser Phe Phe Asn Asp Ala Glu Glu
        115                 120                 125

Ser Lys Ala Thr Ile Lys Asp Tyr Ala Asp Ser Asn Cys Leu Lys Thr
    130                 135                 140

Ser Gly Ile Lys Ser Gln Asp Cys His Ser His Ser Arg Thr Ser Leu
145                 150                 155                 160

Gln Ser Ser His Leu Trp Glu Phe Val Arg Asp Leu Leu Leu Ser Pro
                165                 170                 175

Glu Glu Asn Cys Gly Ile Leu Glu Trp Glu Asp Arg Glu Gln Gly Ile
            180                 185                 190

Phe Arg Val Val Lys Ser Glu Ala Leu Ala Lys Met Trp Gly Gln Arg
        195                 200                 205

Lys Lys Asn Asp Arg Met Thr Tyr Glu Lys Leu Ser Arg Ala Leu Arg
    210                 215                 220

Tyr Tyr Tyr Lys Thr Gly Ile Leu Glu Arg Val Asp Arg Arg Leu Val
225                 230                 235                 240

Tyr Lys Phe Gly Lys Asn Ala His Gly Trp Gln Glu Asp Lys Leu
                245                 250                 255
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1415 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 134..614

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
TGTCTGTAGG TGTCACTTAT ATCACAAGGC TACAGGTGTC TTTATTTCCA CTGCACGCTG      60

GTGCTGGGAG CGCCTGCCTT CTCTTGCCTT GAAAGCCTCC TCTTTGGACC TAGCCACCGC     120
```

```
TGCCCTCACG GTA ATG TTG GAC TCG GTG ACA CAC AGC ACC TTC CTG CCT          169
               Met Leu Asp Ser Val Thr His Ser Thr Phe Leu Pro
                 1               5                  10

AAT GCA TCC CTC TGC GAT CCC CTG ATG TCG TGG ACT GAT CTG TTC AGC          217
Asn Ala Ser Leu Cys Asp Pro Leu Met Ser Trp Thr Asp Leu Phe Ser
         15                  20                  25

AAT GAA GAG TAC TAC CCT GCC TTT GAG CAT CAG ACA GAT GCT GAT TCC          265
Asn Glu Glu Tyr Tyr Pro Ala Phe Glu His Gln Thr Asp Ala Asp Ser
     30                  35                  40

AAC TGC TTG AAA ACA AGT GGC ATC AAA AGC CAA GAC TGT CAC AGT CAT          313
Asn Cys Leu Lys Thr Ser Gly Ile Lys Ser Gln Asp Cys His Ser His
 45                  50                  55                  60

AGT AGA ACA AGC CTC CAA AGT TCT CAT CTA TGG GAA TTT GTA CGA GAC          361
Ser Arg Thr Ser Leu Gln Ser Ser His Leu Trp Glu Phe Val Arg Asp
                 65                  70                  75

CTG CTT CTA TCT CCT GAA GAA AAC TGT GGC ATT CTG GAA TGG GAA GAT          409
Leu Leu Leu Ser Pro Glu Glu Asn Cys Gly Ile Leu Glu Trp Glu Asp
             80                  85                  90

AGG GAA CAA GGA ATT TTT CGG GTG GTT AAA TCG GAA GCC CTG GCA AAG          457
Arg Glu Gln Gly Ile Phe Arg Val Val Lys Ser Glu Ala Leu Ala Lys
         95                  100                 105

ATG TGG GGA CAA AGG AAG AAA AAT GAC AGA ATG ACA TAT GAA AAG TTG          505
Met Trp Gly Gln Arg Lys Lys Asn Asp Arg Met Thr Tyr Glu Lys Leu
     110                 115                 120

AGC AGA GCC CTG AGA TAC TAC TAT AAA ACA GGA ATT TTG GAG CGG GTT          553
Ser Arg Ala Leu Arg Tyr Tyr Tyr Lys Thr Gly Ile Leu Glu Arg Val
125                 130                 135                 140

GAC CGA AGG TTA GTG TAC AAA TTT GGA AAA AAT GCA CAC GGG TGG CAG          601
Asp Arg Arg Leu Val Tyr Lys Phe Gly Lys Asn Ala His Gly Trp Gln
                145                 150                 155

GAA GAC AAG CTA T GATCTGCTCC AGGCATCAAG CTCATTTTAT GGATTTCTGT            654
Glu Asp Lys Leu
            160

CTTTTAAAAC AATCAGATTG CAATAGACAT TCGAAAGGCT TCATTTTCTT CTCTTTTTTT      714

TTAACCTGCA AACATGCTGA TAAAATTTCT CCACATCTCA GCTTACATTT GGATTCAGAG      774

TTGTTGTCTA CGGAGGGTGA GAGCAGAAAC TCTTAAGAAA TCCTTTCTTC TCCCTAAGGG     834

GATGAGGGGA TGATCTTTTG TGGTGTCTTG ATCAAACTTT ATTTTCCTAG AGTTGTGGAA     894

TGACAACAGC CCATGCCATT GATGCTGATC AGAGAAAAAC TATTCAATTC TGCCATTAGA     954

GACACATCCA ATGCTCCCAT CCCAAAGGTT CAAAAGTTTT CAAATAACTG TGGCAGCTCA    1014

CCAAAGGTGG GGGAAAGCAT GATTAGTTTG CAGGTTATGG TAGGAGAGGG TGAGATATAA    1074

GACATACATA CTTTAGATTT TAAATTATTA AAGTCAAAAA TCCATAGAAA AGTATCCCTT    1134

TTTTTTTTTT TGAGACGGGT TCTCACTATG TTGCCCAGGG CTGGTCTTGA ACTCCTATGC    1194

TCAAGTGATC CTCCCACCTC GGCCTCCCAA AGTACTGTGA TTACAAGCGT GAGCCACGGC    1254

ACCTGGGCAG AAAAGTATCT TAATTAATGA AAGAGCTAAG CCATCAAGCT GGGACTTAAT    1314
```

TGGATTTAAC ATAGGTTCAC AGAAAGTTTC CTAACCAGAG CATCTTTTTG ACCACTCAGC 1374

AAAACTTCCA CAGACATCCT TCTGGACTTA AACCGGAATT C 1415

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 160 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Leu Asp Ser Val Thr His Ser Thr Phe Leu Pro Asn Ala Ser Leu
 1               5                  10                  15

Cys Asp Pro Leu Met Ser Trp Thr Asp Leu Phe Ser Asn Glu Glu Tyr
            20                  25                  30

Tyr Pro Ala Phe Glu His Gln Thr Asp Ala Asp Ser Asn Cys Leu Lys
        35                  40                  45

Thr Ser Gly Ile Lys Ser Gln Asp Cys His Ser His Ser Arg Thr Ser
    50                  55                  60

Leu Gln Ser Ser His Leu Trp Glu Phe Val Arg Asp Leu Leu Leu Ser
65                  70                  75                  80

Pro Glu Glu Asn Cys Gly Ile Leu Glu Trp Glu Asp Arg Glu Gln Gly
                85                  90                  95

Ile Phe Arg Val Val Lys Ser Glu Ala Leu Ala Lys Met Trp Gly Gln
            100                 105                 110

Arg Lys Lys Asn Asp Arg Met Thr Tyr Glu Lys Leu Ser Arg Ala Leu
        115                 120                 125

Arg Tyr Tyr Tyr Lys Thr Gly Ile Leu Glu Arg Val Asp Arg Arg Leu
    130                 135                 140

Val Tyr Lys Phe Gly Lys Asn Ala His Gly Trp Gln Glu Asp Lys Leu
145                 150                 155                 160
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2224 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 117..876

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATCACACGGC TACAGGTGCC TTTATTTCTA CAGTCCGCTG GTGCTGGGAG CGCGCTTGCC 60

```
TTCTCTTGCC TTGAAAGCCT TCTGTCTGGA CCTAGCCACC ACTTGTCTTC ACGGTG          116

ATG TTG GAC TCC GTA ACC CAT AGC ACC TTC CTG CCC AAC GCA TCC TTC        164
Met Leu Asp Ser Val Thr His Ser Thr Phe Leu Pro Asn Ala Ser Phe
  1               5                  10                  15

TGT GAC CCC CTG ATG CCT TGG ACC GAT CTG TTC AGC AAT GAA GAC TAC        212
Cys Asp Pro Leu Met Pro Trp Thr Asp Leu Phe Ser Asn Glu Asp Tyr
            20                  25                  30

TAC CCT GCC TTT GAG CAT CAG ACA GCC TGT GAT TCC TAC TGG ACA TCA        260
Tyr Pro Ala Phe Glu His Gln Thr Ala Cys Asp Ser Tyr Trp Thr Ser
        35                  40                  45

GTG CAC CCT GAA TAC TGG ACC AAG CGC CAT GTC TGG GAA TGG CTC CAA        308
Val His Pro Glu Tyr Trp Thr Lys Arg His Val Trp Glu Trp Leu Gln
    50                  55                  60

TTC TGC TGT GAC CAG TAC AAG CTT GAT GCC AAC TGC ATC TCC TTC TGT        356
Phe Cys Cys Asp Gln Tyr Lys Leu Asp Ala Asn Cys Ile Ser Phe Cys
65                  70                  75                  80

CAC TTC AAC ATC AGC GGC CTG CAG CTC TGC AGC ATG ACG CAG GAG GAG        404
His Phe Asn Ile Ser Gly Leu Gln Leu Cys Ser Met Thr Gln Glu Glu
                85                  90                  95

TTC ATT GAG GCA GCC GGC ATC TGT GGG GAG TAC CTG TAC TTC ATT CTC        452
Phe Ile Glu Ala Ala Gly Ile Cys Gly Glu Tyr Leu Tyr Phe Ile Leu
            100                 105                 110

CAG AAC ATT CGC TCG CAA GGT TAC TCC TTT TTC AAT GAT GCT GAA GAG        500
Gln Asn Ile Arg Ser Gln Gly Tyr Ser Phe Phe Asn Asp Ala Glu Glu
        115                 120                 125

ACC AAG ACT GGC ATC AAA GAC TAT GCT GAT TCC AGT TGC TTG AAA ACA        548
Thr Lys Thr Gly Ile Lys Asp Tyr Ala Asp Ser Ser Cys Leu Lys Thr
    130                 135                 140

AGT GGC ATC AAG AGT CAA GAC TGT CAC AGC CGA ACA AGC CTC CAA AGT        596
Ser Gly Ile Lys Ser Gln Asp Cys His Ser Arg Thr Ser Leu Gln Ser
145                 150                 155                 160

TCT CAC CTG TGG GAA TTT GTC AGA GAC TTG CTG CTG TCC CCT GAA GAG        644
Ser His Leu Trp Glu Phe Val Arg Asp Leu Leu Leu Ser Pro Glu Glu
                165                 170                 175

AAC TGT GGC ATC CTG GAA TGG GAA GAC AGG GAG CAG GGC ATT TTC CGA        692
Asn Cys Gly Ile Leu Glu Trp Glu Asp Arg Glu Gln Gly Ile Phe Arg
            180                 185                 190

GTG GTT AAG TCA GAA GCC CTG GCA AAG ATG TGG GGA CAA AGG AAG AAG        740
Val Val Lys Ser Glu Ala Leu Ala Lys Met Trp Gly Gln Arg Lys Lys
        195                 200                 205

AAT GAC AGG ATG ACG TAC GAG AAG CTG AGC CGA GCC CTG AGA TAC TAC        788
Asn Asp Arg Met Thr Tyr Glu Lys Leu Ser Arg Ala Leu Arg Tyr Tyr
    210                 215                 220

TAT AAA ACG AGA ATT CTG GAG CGG GTT GAC CGG AGG TTA GTG TAC AAA        836
Tyr Lys Thr Arg Ile Leu Glu Arg Val Asp Arg Arg Leu Val Tyr Lys
225                 230                 235                 240

TTT GGA AAG AAC GCG CAC GGG TGG CAG GAA GAG AAA CTC T GATGGACACC       886
Phe Gly Lys Asn Ala His Gly Trp Gln Glu Glu Lys Leu
                245                 250
```

GGACACCAGG CTCATTTGAT GGATTTCTGT TGTTGGAAAC AATCAGATCA AACTAGACAT 946

TTGAAAGTCT CCCTCCTCCT CCTCCTCCCC CTCCTTCCCC TCCTCTTCTT CCTCCCCCTC 1006

CTCCTCTTCA AAACCTACAA ACACACTGAT AAAATTTCTG CATGTCTCAG CTTACATTTG 1066

AATTCAGTTG TTGTCTATTG GGGCGATGCC ATCAGCCCTT AAGCAATCGT CTTCATCCCA 1126

AGGGGGAGGA AGGGATGGTC TTGTGGCAAC TTGGTGTGAC ACTGTCTCCT TAATGAAGTG 1186

TTTGGAGCTA AGGGAGCCAG TGTTATGGGT GCTGTTTCAC AAGAGGACCC GTTGCACCAT 1246

TAAGACACAT GATCCTCCCG TTCCAGGGGT TCTGAGCGGT CGACTGAGGC AGCTTGCCTG 1306

TGGTTAGTTT TTAGGAAAGG GAGATGTAAG ACTTCCTTGC TTTAGATTTG AAATTATCAC 1366

AGTTATATTC CATAGAAGAA TTTTTAATTA AAAAAATTTT AGTGGCTAAG CCACTAAACT 1426

GGGACCTAAT TGGATGTAGC CTAAGTTACT AATAAGTTCT TAACCAGATC ACCATTTCCA 1486

ACCACTTAGC CACAGTCACT TGATCCACGG CCAATCCTTC TGAACTTAAC ATCCTTGTAG 1546

TTAGTCACCT TGGGAATTGC TACCTAGATT GTTACCCCCT TCACCTCACT GGTGGCTATC 1606

ATCAGGTCTA CAGTGACCTG ATCAACAGAC ATGTGCATTA ATTTCTAAAT CACTGCTGTG 1666

CCTATGATTC AAACCGTCAG CGTGTTCAGT TTATTGATTC TCTCTGAGGT CGGAATTTAT 1726

TGATTCTCTC TGAGGCTAAG ACATTAAACC TTTACCAAGC AGAGAACGTC CTAACAAGCC 1786

ACGATAGCCG AACACAGCAT CGATCTCTTC TCTTTTCTGA TGAATACTCA AACTTTCCAA 1846

CATATTCTCT TCACAAAAGT AAAGACAGTG AATTTACATC AATCAACGTT CATGGGTTAA 1906

AGTCTGCACT GACATTTCCT TGTCTGCCGT TGCATGCCGT TGGCATGCAA GGTGTTAATG 1966

ACCTGCAACA TGGTGGAGTG CCCTGAACCC TAACTTCCCC AGAGTTGGGA CTGTCTAGTG 2026

ACCGGCACTG AATAGCAATG CAGGCTGAAG ACCTCCAGGT TTAGAATTTA ACCTCAAAAG 2086

TAACTTGTTT TTAAAAAGAA ATGTGAATTA CTGTAAAATA ATCTATTTTT GGATTCGTGT 2146

GTTTTTCAGG TGGATATAGT TTATAAACAA TGTGAATAAA AAATATTTAA CATGTTTAAA 2206

AAAAAAAAAA AAAAAAAA 2224

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1528 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 117..876

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATCACACGGC TACAGGTGCC TTTATTTCTA CAGTCCGCTG GTGCTGGGAG CGCGCTTGCC 60

```
TTCTCTTGCC TTGAAAGCCT TCTGTCTGGA CCTAGCCACC ACTTGTCTTC ACGGTG          116

ATG TTG GAC TCC GTA ACC CAT AGC ACC TTC CTG CCC AAC GCA TCC TTC        164
Met Leu Asp Ser Val Thr His Ser Thr Phe Leu Pro Asn Ala Ser Phe
  1               5                  10                  15

TGT GAC CCC CTG ATG CCT TGG ACC GAT CTG TTC AGC AAT GAA GAC TAC        212
Cys Asp Pro Leu Met Pro Trp Thr Asp Leu Phe Ser Asn Glu Asp Tyr
            20                  25                  30

TAC CCT GCC TTT GAG CAT CAG ACA GCC TGT GAT TCC TAC TGG ACA TCA        260
Tyr Pro Ala Phe Glu His Gln Thr Ala Cys Asp Ser Tyr Trp Thr Ser
        35                  40                  45

GTG CAC CCT GAA TAC TGG ACC AAG CGC CAT GTC TGG GAA TGG CTC CAA        308
Val His Pro Glu Tyr Trp Thr Lys Arg His Val Trp Glu Trp Leu Gln
    50                  55                  60

TTC TGC TGT GAC CAG TAC AAG CTT GAT GCC AAC TGC ATC TCC TTC TGT        356
Phe Cys Cys Asp Gln Tyr Lys Leu Asp Ala Asn Cys Ile Ser Phe Cys
65                  70                  75                  80

CAC TTC AAC ATC AGC GGC CTG CAG CTC TGC AGC ATG ACG CAG GAG GAG        404
His Phe Asn Ile Ser Gly Leu Gln Leu Cys Ser Met Thr Gln Glu Glu
                85                  90                  95

TTC ATT GAG GCA GCC GGC ATC TGT GGG GAG TAC CTG TAC TTC ATT CTC        452
Phe Ile Glu Ala Ala Gly Ile Cys Gly Glu Tyr Leu Tyr Phe Ile Leu
            100                 105                 110

CAG AAC ATT CGC TCG CAA GGT TAC TCC TTT TTC AAT GAT GCT GAA GAG        500
Gln Asn Ile Arg Ser Gln Gly Tyr Ser Phe Phe Asn Asp Ala Glu Glu
        115                 120                 125

ACC AAG ACT GGC ATC AAA GAC TAT GCT GAT TCC AGT TGC TTG AAA ACA        548
Thr Lys Thr Gly Ile Lys Asp Tyr Ala Asp Ser Ser Cys Leu Lys Thr
    130                 135                 140

AGT GGC ATC AAG AGT CAA GAC TGT CAC AGC CGA ACA AGC CTC CAA AGT        596
Ser Gly Ile Lys Ser Gln Asp Cys His Ser Arg Thr Ser Leu Gln Ser
145                 150                 155                 160

TCT CAC CTG TGG GAA TTT GTC AGA GAC TTG CTG CTG TCC CCT GAA GAG        644
Ser His Leu Trp Glu Phe Val Arg Asp Leu Leu Leu Ser Pro Glu Glu
                165                 170                 175

AAC TGT GGC ATC CTG GAA TGG GAA GAC AGG GAG CAG GGC ATT TTC CGA        692
Asn Cys Gly Ile Leu Glu Trp Glu Asp Arg Glu Gln Gly Ile Phe Arg
            180                 185                 190

GTG GTT AAG TCA GAA GCC CTG GCA AAG ATG TGG GGA CAA AGG AAG AAG        740
Val Val Lys Ser Glu Ala Leu Ala Lys Met Trp Gly Gln Arg Lys Lys
        195                 200                 205

AAT GAC AGG ATG ACG TAC GAG AAG CTG AGC CGA GCC CTG AGA TAC TAC        788
Asn Asp Arg Met Thr Tyr Glu Lys Leu Ser Arg Ala Leu Arg Tyr Tyr
    210                 215                 220

TAT AAA ACG AGA ATT CTG GAG CGG GTT GAC CGG AGG TTA GTG TAC AAA        836
Tyr Lys Thr Arg Ile Leu Glu Arg Val Asp Arg Arg Leu Val Tyr Lys
225                 230                 235                 240

TTT GGA AAG AAC GCG CAC GGG TGG CAG GAA GAG AAA CTC T GATGGACACC       886
Phe Gly Lys Asn Ala His Gly Trp Gln Glu Glu Lys Leu
                245                 250
```

GGACACCAGG CTCATTTGAT GGATTTCTGT TGTTGGAAAC AATCAGATCA AACTAGACAT 946

TTGAAAGTCT CCCTCCTCCT CCTCCTCCCC CTCCTTCCCC TCCTCTTCTT CCTCCCCCTC 1006

CTCCTCTTCA AAACCTACAA ACACACTGAT AAAATTTCTG CATGTCTCAG CTTACATTTG 1066

AATTCAGTTG TTGTCTATTG GGGCGATGCC ATCAGCCCTT AAGCAATCGT CTTCATCCCA 1126

AGGGGGAGGA AGGGATGGTC TTGTGGCAAC TTGGTGTGAC ACTGTCTCCT TAATGAAGTG 1186

TTTGGAGCTA AGGGAGCCAG TGTTATGGGT GCTGTTTCAC AAGAGGACCC GTTGCACCAT 1246

TAAGACACAT GATCCTCCCG TTCCAGGGGT TCTGAGCGGT CGACTGAGGC AGCTTGCCTG 1306

TGGTTAGTTT TTAGGAAAGG GAGATGTAAG ACTTCCTTGC TTTAGATTTG AAATTATCAC 1366

AGTTATATTC CATAGAAGAA TTTTTAATTA AAAAAATTTT AGTGGCTAAG CCACTAAACT 1426

GGGACCTAAT TGGATGTAGC CTAAGTTACT AATAAGTTCT TAACCAGATC ACCATTTCCA 1486

ACCACTTAGC CACAGTCACT TGATCCACGG CCAATCCTTC TG 1528

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 253 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
Met Leu Asp Ser Val Thr His Ser Thr Phe Leu Pro Asn Ala Ser Phe
 1               5                  10                  15

Cys Asp Pro Leu Met Pro Trp Thr Asp Leu Phe Ser Asn Glu Asp Tyr
            20                  25                  30

Tyr Pro Ala Phe Glu His Gln Thr Ala Cys Asp Ser Tyr Trp Thr Ser
        35                  40                  45

Val His Pro Glu Tyr Trp Thr Lys Arg His Val Trp Glu Trp Leu Gln
    50                  55                  60

Phe Cys Cys Asp Gln Tyr Lys Leu Asp Ala Asn Cys Ile Ser Phe Cys
65                  70                  75                  80

His Phe Asn Ile Ser Gly Leu Gln Leu Cys Ser Met Thr Gln Glu Glu
                85                  90                  95

Phe Ile Glu Ala Ala Gly Ile Cys Gly Glu Tyr Leu Tyr Phe Ile Leu
            100                 105                 110

Gln Asn Ile Arg Ser Gln Gly Tyr Ser Phe Phe Asn Asp Ala Glu Glu
        115                 120                 125

Thr Lys Thr Gly Ile Lys Asp Tyr Ala Asp Ser Ser Cys Leu Lys Thr
    130                 135                 140

Ser Gly Ile Lys Ser Gln Asp Cys His Ser Arg Thr Ser Leu Gln Ser
145                 150                 155                 160
```

Ser His Leu Trp Glu Phe Val Arg Asp Leu Leu Leu Ser Pro Glu Glu
165 170 175

Asn Cys Gly Ile Leu Glu Trp Glu Asp Arg Glu Gln Gly Ile Phe Arg
180 185 190

Val Val Lys Ser Glu Ala Leu Ala Lys Met Trp Gly Gln Arg Lys Lys
195 200 205

Asn Asp Arg Met Thr Tyr Glu Lys Leu Ser Arg Ala Leu Arg Tyr Tyr
210 215 220

Tyr Lys Thr Arg Ile Leu Glu Arg Val Asp Arg Arg Leu Val Tyr Lys
225 230 235 240

Phe Gly Lys Asn Ala His Gly Trp Gln Glu Glu Lys Leu
245 250

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCCAGTCTTG GTCTCTTCAG CATC 24

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGGAGATGCA GTTGGCATCA AGCT 24

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGCCAGTGTT ATGGGTGCTG 20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACAGTCACTT GATCCACGGC CAATCC 26

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12

GGGTGGCAGG AAGACAAGCT ATGA 24

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CCAATTAAGT CCCAGCTTGA TGGC 24

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CGGGATCCTT GGACTCCGTA ACCCATAGC 29

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCAGATCTCA GAGTTTCTCT TCCTGCC 27

```
    atcacacggctacaggtgcctttatttctacagtccgctggtgctgggagcgcgct      56
tgccttctcttgccttgaaagccttctgtctggac                            91
                                    ctagccaccacttgtcttcacggtg     116

ATGTTGGACTCCGTAACCCATAGCACCTTCCTGCCCAACGCATCCTTCTGTGACCCCCTG     176

ATGCCTTGGACCGATCTGTTCAGCAATGAAGACTACTACCCTGCCTTTGAGCATCAGACA     236

GCCTGTGATTCCTACTGGACATCAGTGCACCCTGAATACTGGACCAAGCGCCATGTCTGG     296

GAATGGCTCCAATTCTGCTGTGACCAGTACAAGCTTGATGCCAACTGCATCTCCTTCTGT     356

CACTTCAACATCAGCGGCCTGCAGCTCTGCAGCATGACGCAGGAGGAGTTCATTGAGGCA     416

GCCGGCATCTGTGGGGAGTACCTGTACTTCATTCTCCAGAACATTCGCTCGCAAGGTTAC     476

TCCTTTTTCAATGATGCTGAAGAGACCAAGACTGGCATCAAAGACTATGCTGATTCCAGT     536

TGCTTGAAAACAAGTGGCATCAAGAGTCAAGACTGTCACAGCCGAACAAGCCTCCAAAGT     596

TCTCACCTGTGGGAATTTGTCAGAGACTTGCTGCTGTCCCCTGAAGAGAACTGTGGCATC     656

CTGGAATGGGAAGACAGGGAGCAGGGCATTTTCCGAGTGGTTAAGTCAGAAGCCCTGGCA     716

AAGATGTGGGGACAAAGGAAGAAGAATGACAGGATGACGTACGAGAAGCTGAGCCGAGCC     776

CTGAGATACTACTATAAAACGAGAATTCTGGAGCGGGTTGACCGGAGGTTAGTGTACAAA     836

TTTGGAAAGAACGCGCACGGGTGGCAGGAAGAGAAACTCTGAtggacaccggacaccagg     896

ctcatttgatggatttctgttgttggaaacaatcagatcaaactagacatttgaaagtct      956
ccctcctcctcctcctcccccctccttccccctcctcttcttcctcccccctcctcctcttca  1016
aaacctacaaacacactgataaaatttctgcatgtctcagcttacatttgaattcagttg     1076
ttgtctattggggcgatgccatcagcccttaagcaatcgtcttcatcccaaggggggagga   1136
agggatggtcttgtggcaacttggtgtgacactgtctccttaatgaagtgtttggagcta     1196
agggagccagtgttatgggtgctgtttcacaagaggacccgttgcaccattaagacacat     1256
gatcctcccgttccaggggttctgagcggtcgactgaggcagcttgcctgtggttagttt     1316
ttaggaaagggagatgtaagacttccttgctttagatttgaaattatcacagttatattc     1376
catagaagaatttttaattaaaaaaattttagtggctaagccactaaactgggacctaat     1436
tggatgtagcctaagttactaataagttcttaaccagatcaccatttccaaccacttagc     1496
cacagtcacttgatccacggccaatccttctg                                1528
                                    aacttaacatccttgtagttagtcacct 1556
tgggaattgctacctagattgttacccccttcacctcactggtggctatcatcaggtcta    1616
cagtgacctgatcaacagacatgtgcattaatttctaaatcactgctgtgcctatgattc    1676
aaaccgtcagcgtgttcagtttattgattctctctgaggtcggaatttattgattctctc    1736
tgaggctaagacattaaacctttaccaagcagagaacgtcctaacaagccacgatagccg    1796
aacacagcatcgatctcttctcttttctgatgaatactcaaactttccaacatattctct    1856
tcacaaaagtaaagacagtgaatttacatcaatcaacgttcatgggttaaagtctgcact    1916
gacatttccttgtctgccgttgcatgccgttggcatgcaaggtgttaatgacctgcaaca    1976
tggtggagtgccctgaaccctaacttccccagagttgggactgtctagtgaccggcactg    2036
aatagcaatgcaggctgaagacctccaggtttagaatttaacctcaaaagtaacttgttt    2096
ttaaaaagaaatgtgaattactgtaaaataatctatttttggattcgtgtgtttttcagg    2156
tggatatagtttataaacaatgtgaataaaaaatatttaacatgtttaaaaaaaaaaaaa    2216
aaaaaaaa                                                        2224
```

FIGURE 1a

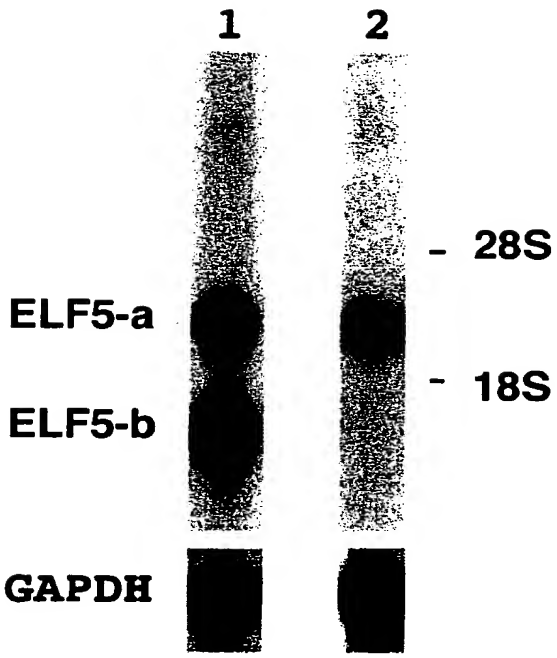


FIGURE 1b

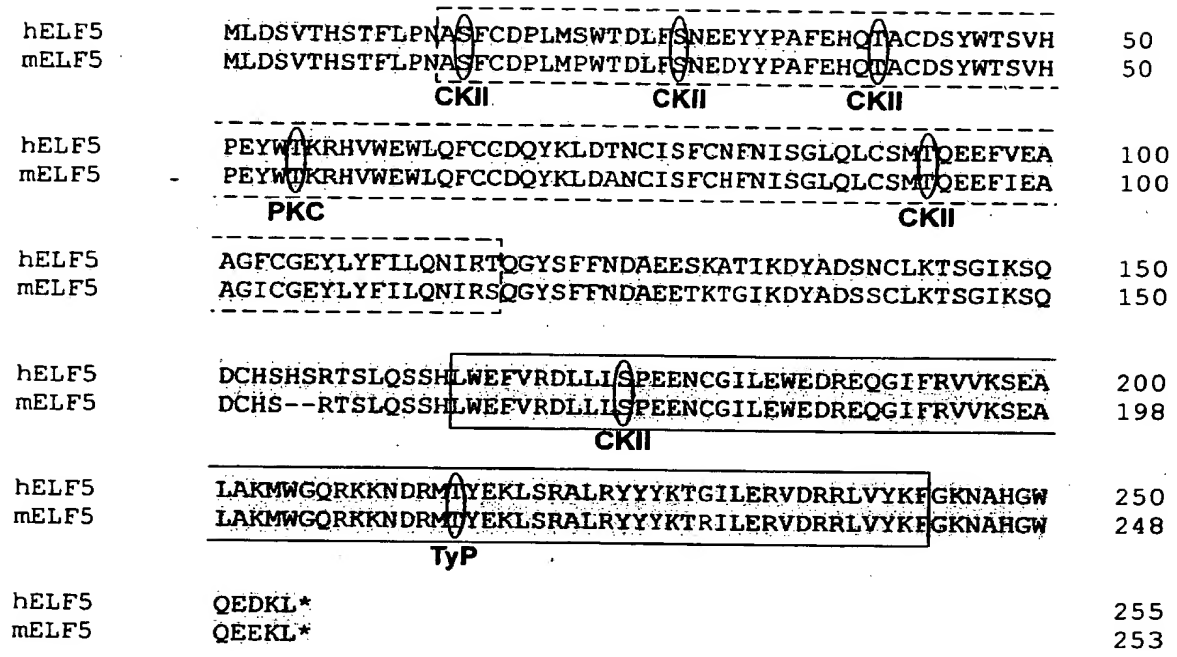
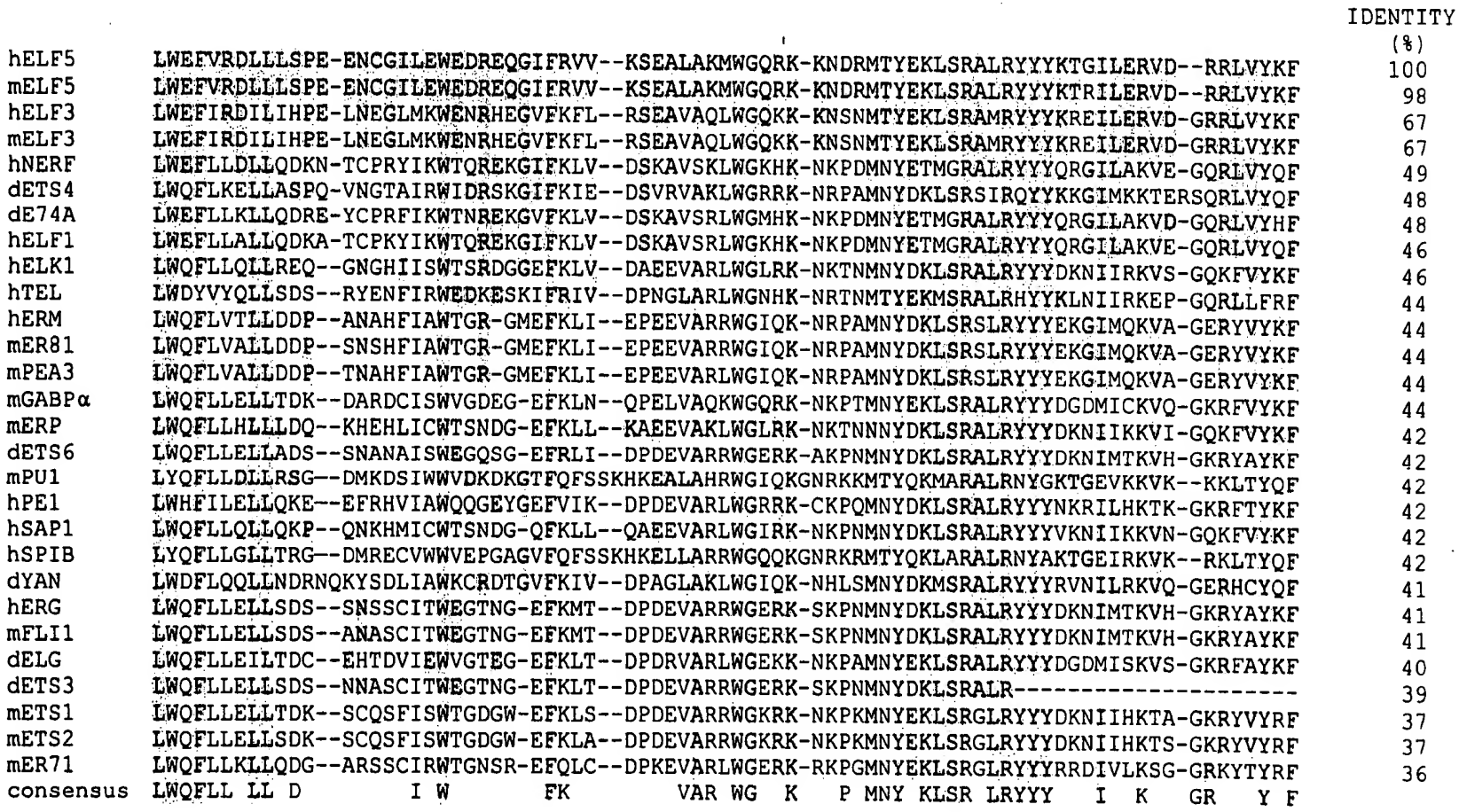

hELF5 MLDSVTHSTFLPNASFCDPLMSWTDLFSNEEYYPAFEHQTACDSYWTSVH 50
mELF5 MLDSVTHSTFLPNASFCDPLMPWTDLFSNEDYYPAFEHQTACDSYWTSVH 50
CKII
CKII
CKII
hELF5 PEYWTKRHVWEWLQFCCDQYKLDTNCISFCNFNISGLQLCSMTQEEFVEA 100
mELF5 PEYWTKRHVWEWLQFCCDQYKLDANCISFCHFNISGLQLCSMTQEEFIEA 100
PKC
CKII
hELF5 AGFCGEYLYFILQNIRTQGYSFFNDAEESKATIKDYADSNCLKTSGIKSQ 150
mELF5 AGICGEYLYFILQNIRSQGYSFFNDAEETKTGIKDYADSSCLKTSGIKSQ 150
hELF5 DCHSHSRTSLQSSHLWEFVRDLLLSPEENCGILEWEDREQGIFRVVKSEA 200
mELF5 DCHS--RTSLQSSHLWEFVRDLLLSPEENCGILEWEDREQGIFRVVKSEA 198
CKII
hELF5 LAKMWGQRKKNDRMIYEKLSRALRYYYKTGILERVDRRLVYKFGKNAHGW 250
mELF5 LAKMWGQRKKNDRMIYEKLSRALRYYYKTRILERVDRRLVYKFGKNAHGW 248
TyP
hELF5 QEDKL* 255
mELF5 QEEKL* 253

FIGURE 2a



		IDENTITY (%)
hELF5	LWEFVRDLLLSPE-ENCGILEWEDREQGIFRVV--KSEALAKMWGQRK-KNDRMTYEKLSRALRYYYKTGILERVD--RRLVYKF	100
mELF5	LWEFVRDLLLSPE-ENCGILEWEDREQGIFRVV--KSEALAKMWGQRK-KNDRMTYEKLSRALRYYYKTRILERVD--RRLVYKF	98
hELF3	LWEFIRDILIHPE-LNEGLMKWENRHEGVFKFL--RSEAVAQLWGQKK-KNSNMTYEKLSRAMRYYYKREILERVD-GRRLVYKF	67
mELF3	LWEFIRDILIHPE-LNEGLMKWENRHEGVFKFL--RSEAVAQLWGQKK-KNSNMTYEKLSRAMRYYYKREILERVD-GRRLVYKF	67
hNERF	LWEFLLDLLQDKN-TCPRYIKWTQREKGIFKLV--DSKAVSKLWGKHK-NKPDMNYETMGRALRYYYQRGILAKVE-GQRLVYQF	49
dETS4	LWQFLKELLASPQ-VNGTAIRWIDRSKGIFKIE--DSVRVAKLWGRRK-NRPAMNYDKLSRSIRQYYKKGIMKKTERSQRLVYQF	48
dE74A	LWEFLLKLLQDRE-YCPRFIKWTNREKGVFKLV--DSKAVSRLWGMHK-NKPDMNYETMGRALRYYYQRGILAKVD-GQRLVYHF	48
hELF1	LWEFLLALLQDKA-TCPKYIKWTQREKGIFKLV--DSKAVSRLWGKHK-NKPDMNYETMGRALRYYYQRGILAKVD-GQRLVYHF	48
hELK1	LWQFLLQLLREQ--GNGHIISWTSRDGGEFKLV--DAEEVARLWGLRK-NKTNMNYDKLSRALRYYYDKNIIRKVS-GQKFVYKF	46
hTEL	LWDYVYQLLSDS--RYENFIRWEDKESKIFRIV--DPNGLARLWGNHK-NRTNMTYEKMSRALRHYYKLNIIRKEP-GQRLLFRF	46
hERM	LWQFLVTLLDDP--ANAHFIAWTGR-GMEFKLI--EPEEVARRWGIQK-NRPAMNYDKLSRSLRYYYEKGIMQKVA-GERYVYKF	44
mER81	LWQFLVALLDDP--SNSHFIAWTGR-GMEFKLI--EPEEVARRWGIQK-NRPAMNYDKLSRSLRYYYEKGIMQKVA-GERYVYKF	44
mPEA3	LWQFLVALLDDP--TNAHFIAWTGR-GMEFKLI--EPEEVARLWGIQK-NRPAMNYDKLSRSLRYYYEKGIMQKVA-GERYVYKF	44
mGABPα	LWQFLLELLTDK--DARDCISWVGDEG-EFKLN--QPELVAQKWGQRK-NKPTMNYEKLSRALRYYYDGDMICKVQ-GKRFVYKF	44
mERP	LWQFLLHLLLDQ--KHEHLICWTSNDG-EFKLL--KAEEVAKLWGLRK-NKTNNNYDKLSRALRYYYDKNIIKKVI-GQKFVYKF	42
dETS6	LWQFLLELLADS--SNANAISWEGQSG-EFRLI--DPDEVARRWGERK-AKPNMNYDKLSRALRYYYDKNIMTKVH-GKRYAYKF	42
mPU1	LYQFLLDLLRSG--DMKDSIWWVDKDKGTFQFSSKHKEALAHRWGIQKGNRKKMTYQKMARALRNYGKTGEVKKVK--KKLTYQF	42
hPE1	LWHFILELLQKE--EFRHVIAWQQGEYGEFVIK--DPDEVARLWGRRK-CKPQMNYDKLSRALRYYYNKRILHKTK-GKRFTYKF	42
hSAP1	LWQFLLQLLQKP--QNKHMICWTSNDG-QFKLL--QAEEVARLWGIRK-NKPNMNYDKLSRALRYYYVKNIIKKVN-GQKFVYKF	42
hSPIB	LYQFLLGLLTRG--DMRECVWWVEPGAGVFQFSSKHKELLARRWGQQKGNRKRMTYQKLARALRNYAKTGEIRKVK--RKLTYQF	42
dYAN	LWDFLQQLLNDRNQKYSDLIAWKCRDTGVFKIV--DPAGLAKLWGIQK-NHLSMNYDKMSRALRYYYRVNILRKVQ-GERHCYQF	41
hERG	LWQFLLELLSDS--SNSSCITWEGTNG-EFKMT--DPDEVARRWGERK-SKPNMNYDKLSRALRYYYDKNIMTKVH-GKRYAYKF	41
mFLI1	LWQFLLELLSDS--ANASCITWEGTNG-EFKMT--DPDEVARRWGERK-SKPNMNYDKLSRALRYYYDKNIMTKVH-GKRYAYKF	41
dELG	LWQFLLEILTDC--EHTDVIEWVGTEG-EFKLT--DPDRVARLWGEKK-NKPAMNYEKLSRALRYYYDGDMISKVS-GKRFAYKF	40
dETS3	LWQFLLELLSDS--NNASCITWEGTNG-EFKLT--DPDEVARRWGERK-SKPNMNYDKLSRALR----------	39
mETS1	LWQFLLELLTDK--SCQSFISWTGDGW-EFKLS--DPDEVARRWGKRK-NKPKMNYEKLSRGLRYYYDKNIIHKTA-GKRYVYRF	37
mETS2	LWQFLLELLSDK--SCQSFISWTGDGW-EFKLA--DPDEVARRWGKRK-NKPKMNYEKLSRGLRYYYDKNIIHKTS-GKRYVYRF	37
mER71	LWQFLLKLLQDG--ARSSCIRWTGNSR-EFQLC--DPKEVARLWGERK-RKPGMNYEKLSRGLRYYYRRDIVLKSG-GRKYTYRF	36
consensus	LWQFLL LL D I W FK VAR WG K P MNY KLSR LRYYY I K GR Y F	

FIGURE 2b

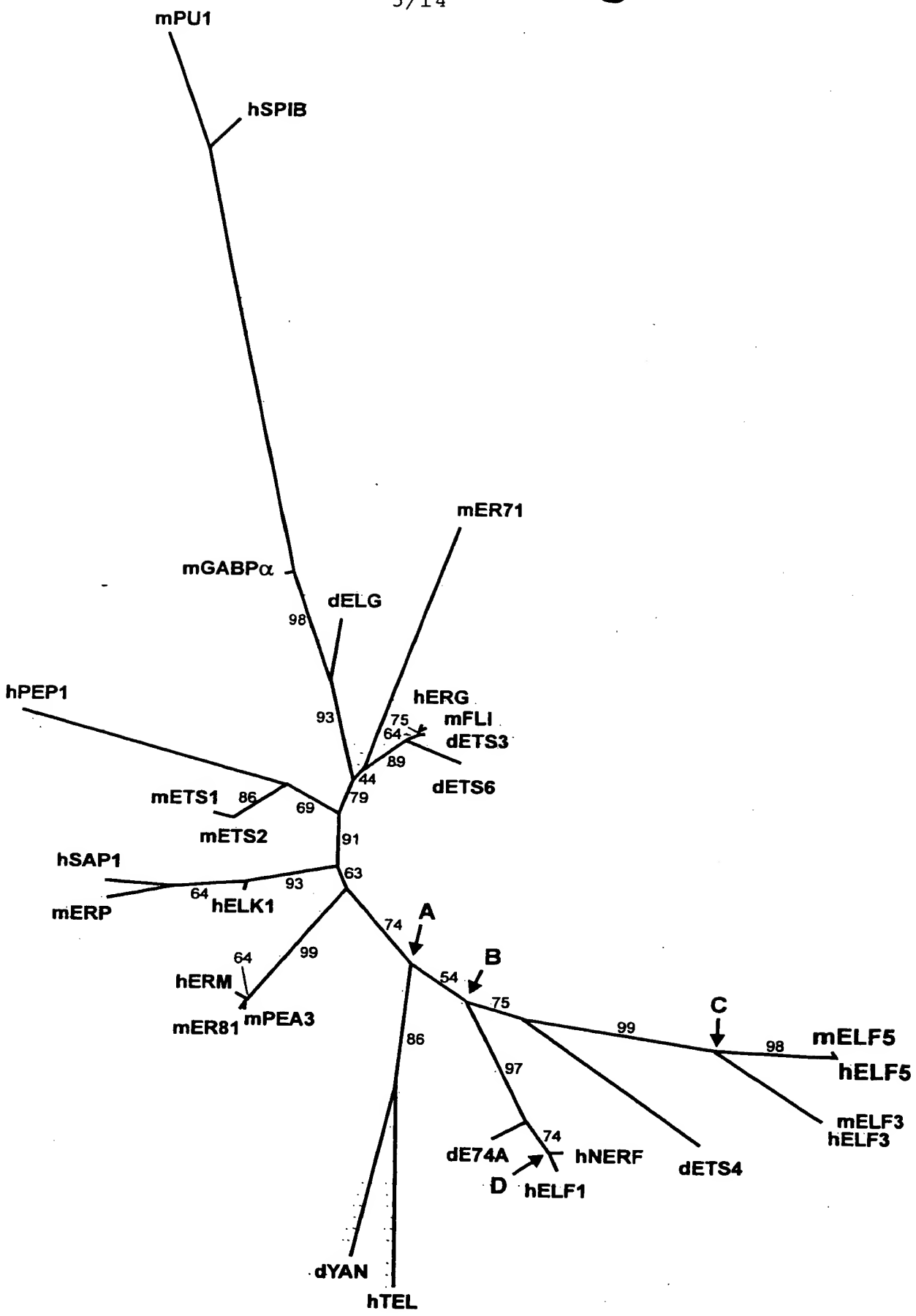

mPU1
hSPIB
mER71
mGABPα
dELG
98
hPEP1
hERG
93
75
mFLI
64
dETS3
89
44
dETS6
mETS1
86
79
69
mETS2
91
hSAP1
63
64
93
hELK1
mERP
74
A
64
99
B
hERM
54
75
C
mER81
mPEA3
99
98
mELF5
86
hELF5
97
mELF3
hELF3
74
dE74A
hNERF
dETS4
D
hELF1
dYAN
hTEL

FIGURE 2c

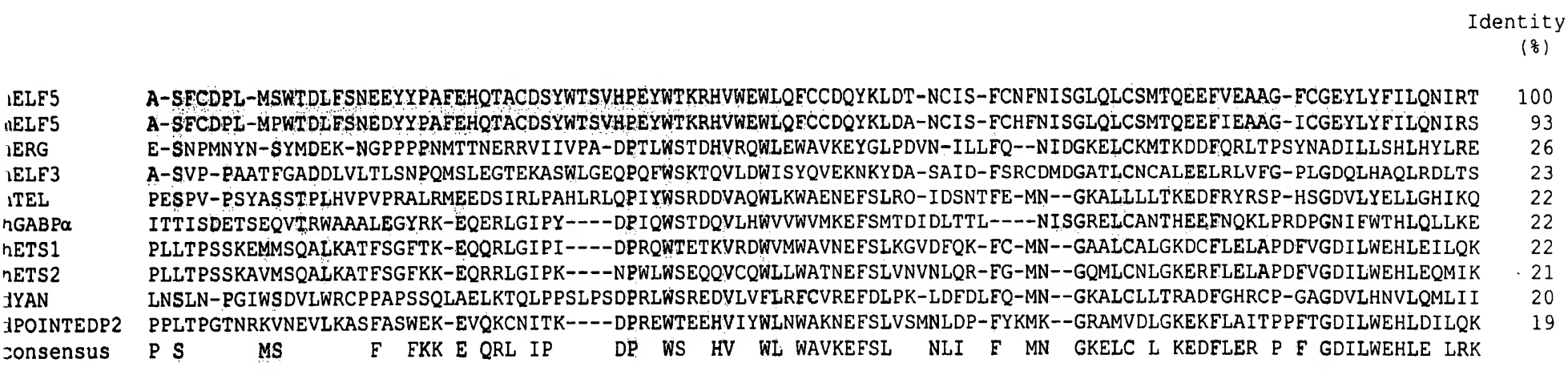

Identity
(%)
ıELF5 A-SFCDPL-MSWTDLFSNEEYYPAFEHQTACDSYWTSVHPEYWTKRHVWEWLQFCCDQYKLDT-NCIS-FCNFNISGLQLCSMTQEEFVEAAG-FCGEYLYFILQNIRT 100
ıELF5 A-SFCDPL-MPWTDLFSNEDYYPAFEHQTACDSYWTSVHPEYWTKRHVWEWLQFCCDQYKLDA-NCIS-FCHFNISGLQLCSMTQEEFIEAAG-ICGEYLYFILQNIRS 93
ıERG E-SNPMNYN-SYMDEK-NGPPPPNMTTNERRVIIVPA-DPTLWSTDHVRQWLEWAVKEYGLPDVN-ILLFQ--NIDGKELCKMTKDDFQRLTPSYNADILLSHLHYLRE 26
ıELF3 A-SVP-PAATFGADDLVLTLSNPQMSLEGTEKASWLGEQPQFWSKTQVLDWISYQVEKNKYDA-SAID-FSRCDMDGATLCNCALEELRLVFG-PLGDQLHAQLRDLTS 23
ıTEL PESPV-PSYASSTPLHVPVPRALRMEEDSIRLPAHLRLQPIYWSRDDVAQWLKWAENEFSLRQ-IDSNTFE-MN--GKALLLLTKEDFRYRSP-HSGDVLYELLGHIKQ 22
ηGABPα ITTISDETSEQVTRWAAALEGYRK-EQERLGIPY----DPIQWSTDQVLHWVVWVMKEFSMTDIDLTTL----NISGRELCANTHEEFNQKLPRDPGNIFWTHLQLLKE 22
ηETS1 PLLTPSSKEMMSQALKATFSGFTK-EQQRLGIPI----DPRQWTETKVRDWVMWAVNEFSLKGVDFQK-FC-MN--GAALCALGKDCFLELAPDFVGDILWEHLEILQK 22
ηETS2 PLLTPSSKAVMSQALKATFSGFKK-EQRRLGIPK----NPWLWSEQQVCQWLLWATNEFSLVNVNLQR-FG-MN--GQMLCNLGKERFLELAPDFVGDILWEHLEQMIK 21
ıYAN LNSLN-PGIWSDVLWRCPPAPSSQLAELKTQLPPSLPSDPRLWSREDVLVFLRFCVREFDLPK-LDFDLFQ-MN--GKALCLLTRADFGHRCP-GAGDVLHNVLQMLII 20
ıPOINTEDP2 PPLTPGTNRKVNEVLKASFASWEK-EVQKCNITK----DPREWTEEHVIYWLNWAKNEFSLVSMNLDP-FYKMK--GRAMVDLGKEKFLAITPPFTGDILWEHLDILQK 19
consensus P S MS F FKK E QRL IP DP WS HV WL WAVKEFSL NLI F MN GKELC L KEDFLER P F GDILWEHLE LRK

FIGURE 2d

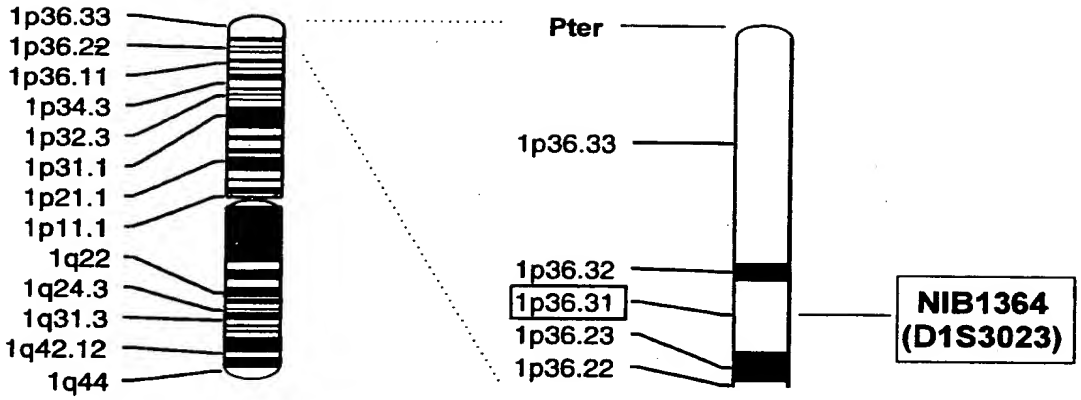

1p36.33
1p36.22
1p36.11
1p34.3
1p32.3
1p31.1
1p21.1
1p11.1
1q22
1q24.3
1q31.3
1q42.12
1q44
Pter
1p36.33
1p36.32
1p36.31
1p36.23
1p36.22
NIB1364
(D1S3023)

FIGURE 3

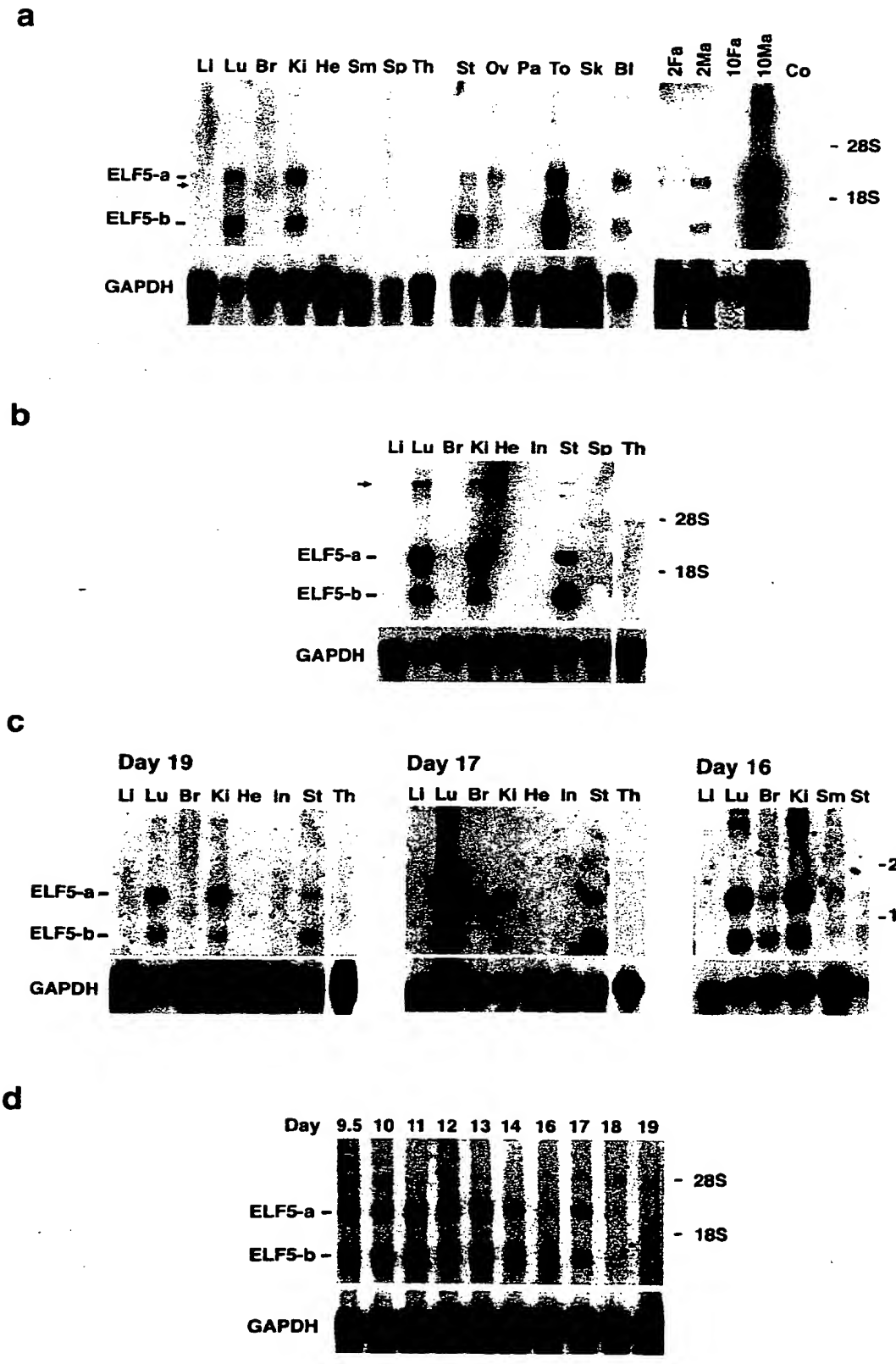
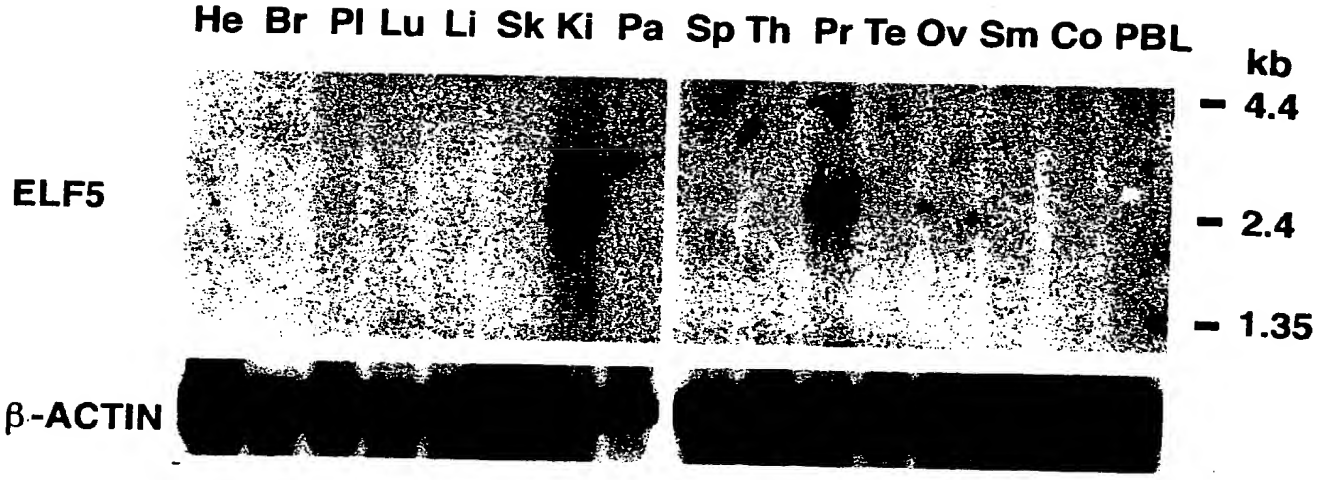

a
Li Lu Br Ki He Sm Sp Th St Ov Pa To Sk Bl 2Fa 2Ma 10Fa 10Ma Co
ELF5-a
ELF5-b
GAPDH
- 28S
- 18S
b
Li Lu Br Ki He In St Sp Th
ELF5-a
ELF5-b
GAPDH
- 28S
- 18S
c
Day 19
Li Lu Br Ki He In St Th
Day 17
Li Lu Br Ki He In St Th
Day 16
Li Lu Br Ki Sm St
ELF5-a
ELF5-b
GAPDH
-28S
-18S
d
Day 9.5 10 11 12 13 14 16 17 18 19
ELF5-a
ELF5-b
GAPDH
- 28S
- 18S

FIGURE 4

a


He Br Pl Lu Li Sk Ki Pa Sp Th Pr Te Ov Sm Co PBL
kb
4.4
2.4
1.35
ELF5
β-ACTIN

b

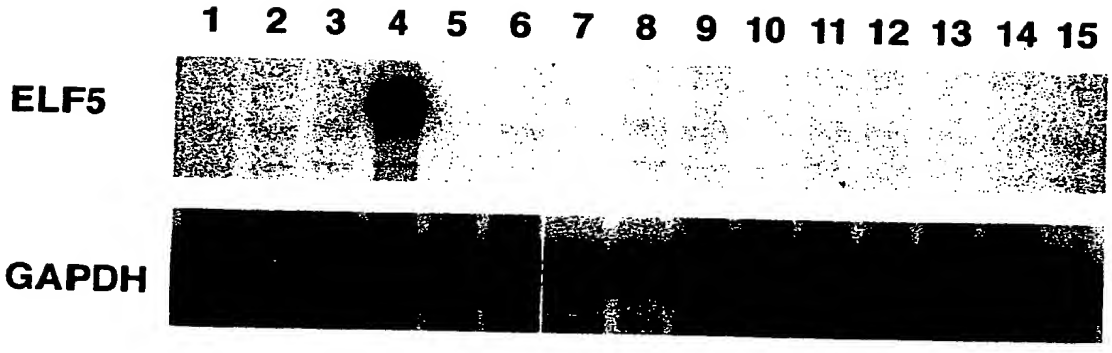

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
ELF5
GAPDH

FIGURE 5

C

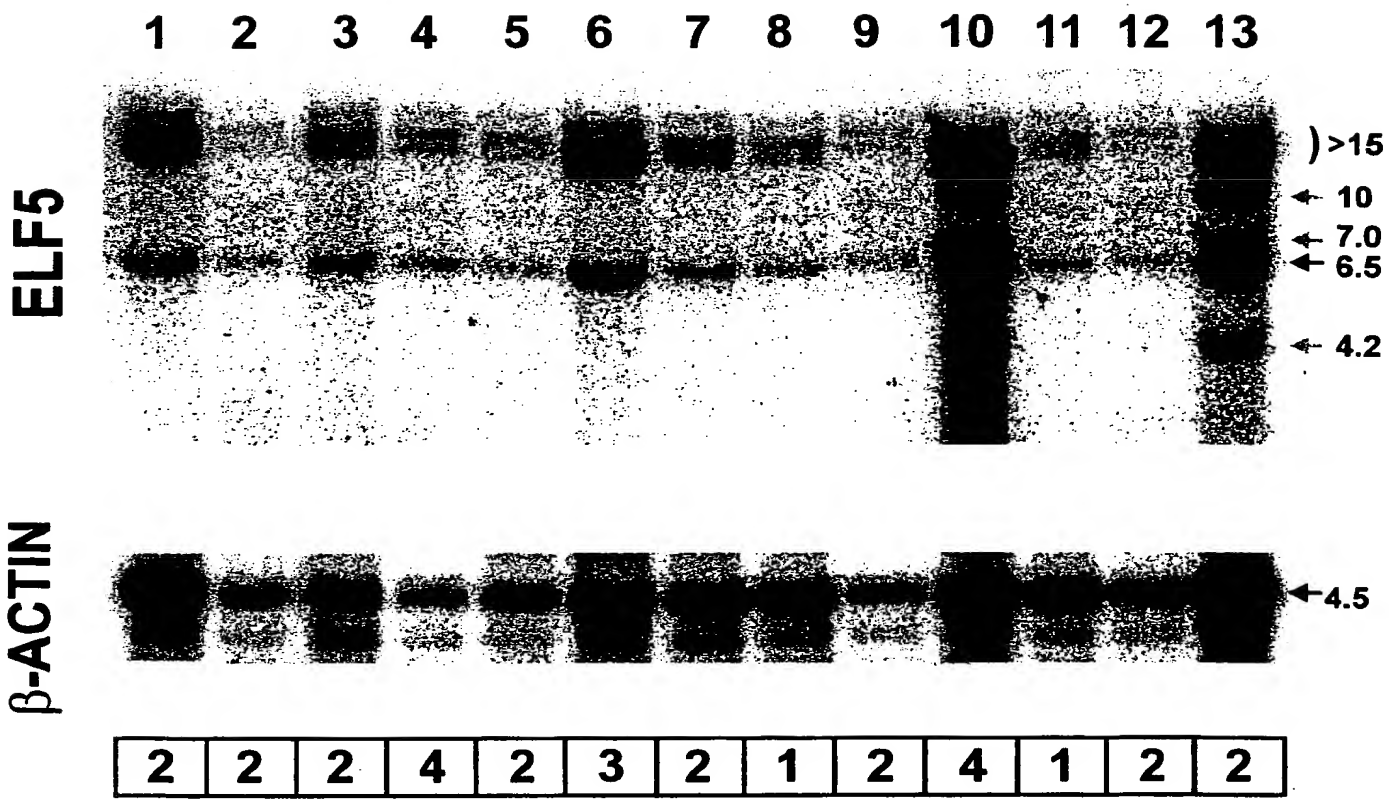

1 2 3 4 5 6 7 8 9 10 11 12 13
ELF5
>15
10
7.0
6.5
4.2
β-ACTIN
4.5
2 2 2 4 2 3 2 1 2 4 1 2 2

FIGURE 5c

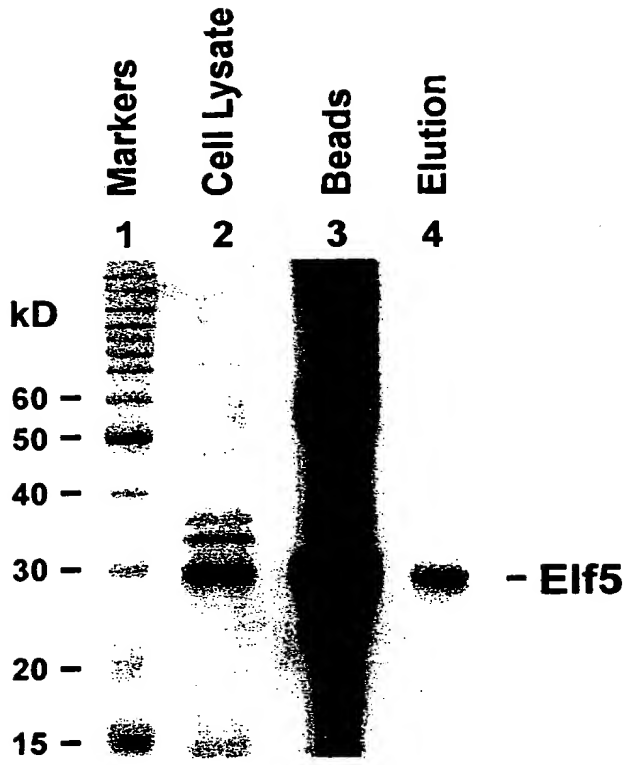

Markers
Cell Lysate
Beads
Elution
1
2
3
4
kD
60
50
40
30
20
15
Elf5

FIGURE 6a

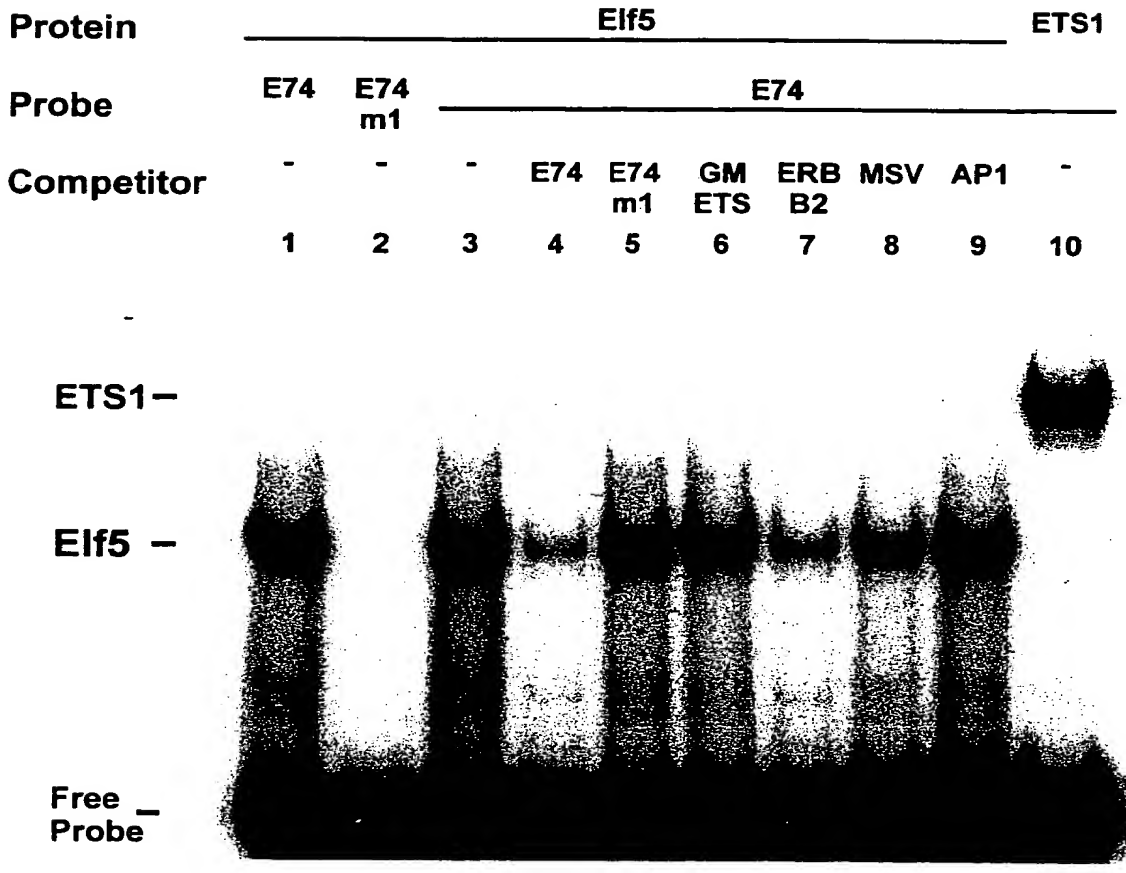

Protein
Elf5
ETS1
Probe
E74
E74 m1
E74
Competitor
-
-
-
E74
E74 m1
GM ETS
ERB B2
MSV
AP1
-
1
2
3
4
5
6
7
8
9
10
ETS1
Elf5
Free Probe

FIGURE 6b

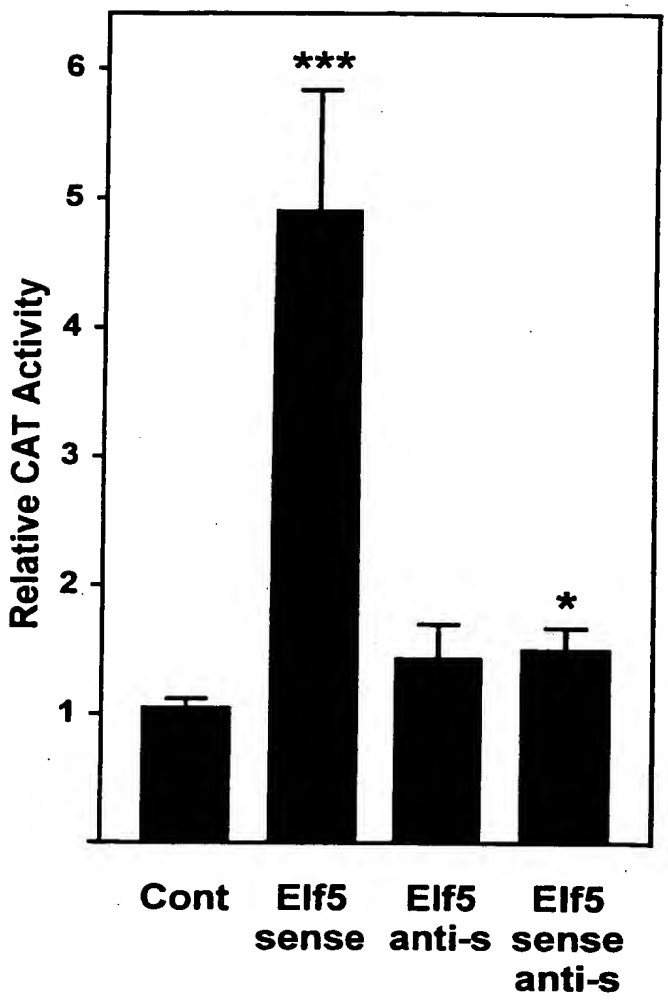

Relative CAT Activity
6
5
4
3
2
1
***
*
Cont
Elf5 sense
Elf5 anti-s
Elf5 sense anti-s

FIGURE 7

ELF5 expression is not detectable in human primary breast cancer cells, but is strongly expressed in adjacent normal epithelium

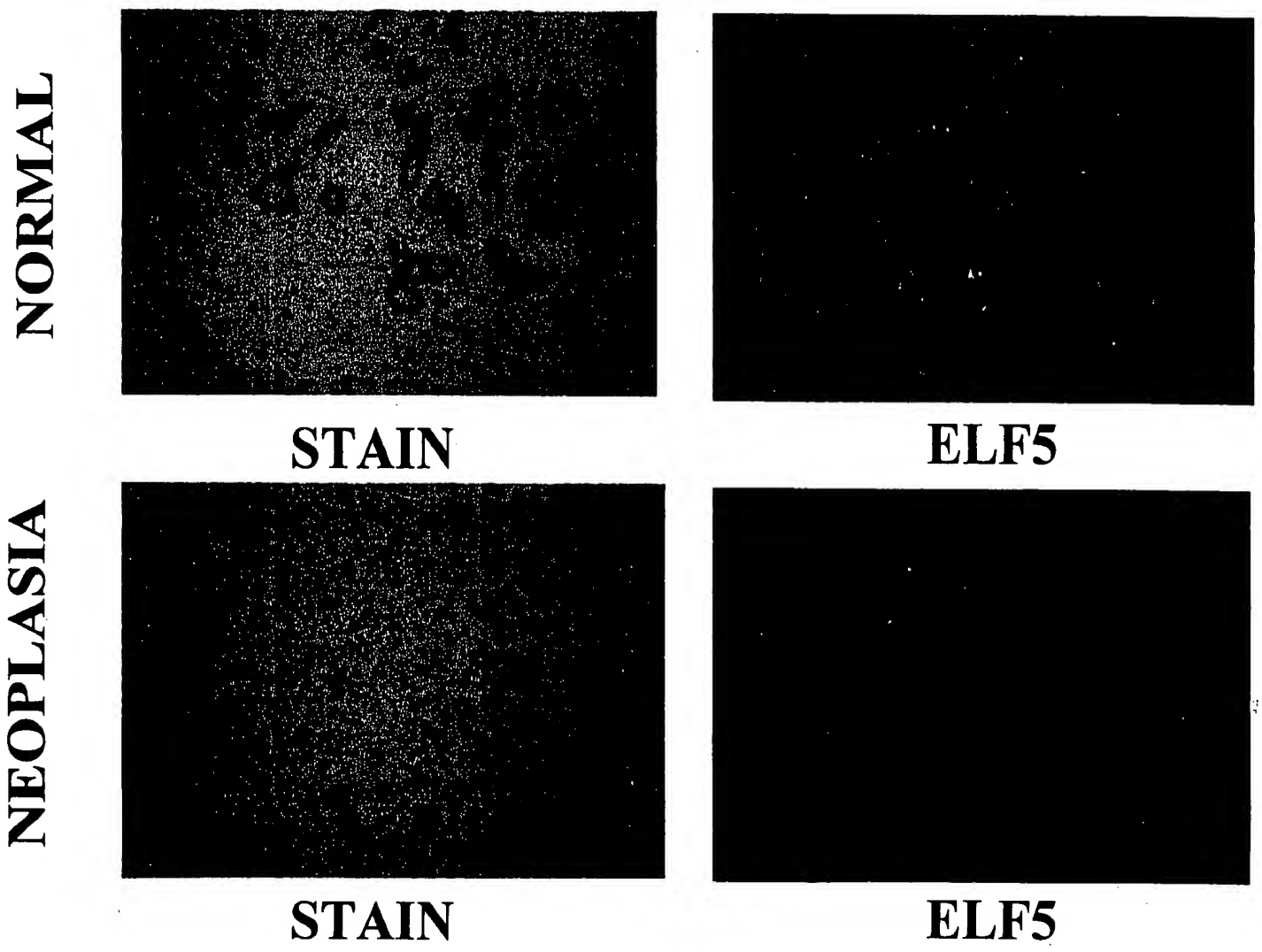


FIGURE 8